Computers, Environment and Urban Systems 75 (2019) 254–263

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Temporal and spatial analysis of the media spotlight

Rafael Prieto Curiel[a,*], Carmen Cabrera Arnau[b], Mara Torres Pinedo[c], Humberto González Ramírez[d], Steven R. Bishop[b]

[a] Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Rd, Oxford OX2 6GG, United Kingdom
[b] Department of Mathematics, University College London, 25 Gower Street, London WC1E 6BT, United Kingdom
[c] Institute for Risk and Disaster Reduction, University College London, 25 Gower Street, London WC1E 6BT, United Kingdom
[d] École Nationale des Travaux Publics de l'État, ENTPE, Universiteé de Lyon 2, 3 rue Maurice Audin, 69518 Vaulx-en-Velin cedex, France

ARTICLE INFO

Keywords:
Scaling
Media coverage

ABSTRACT

Items featured in the news usually have a particular novelty or describe events which result in severe impact. Here, the length of time that a story remains in the media spotlight is investigated as well as the scaling with population size of the amount of attention that the media gives to stories from different cities. Based on Twitter feeds, the media coverage from the major online newspapers in Mexico is analysed over a period either side of a recent powerful earthquake. The amount of coverage given to earthquake-related stories had an initial peak and then exhibited an exponential decay, dropping by half every eight days. Furthermore, the coverage per person usually exhibits a superlinear scaling with population size, so that stories about larger cities are more likely to appear in the news. However, during the immediate post-earthquake weeks, the scaling was no longer superlinear. The observed trends can be interpreted as a fundamental switch in the emergent collective behaviour of media producers and consumers.

1. Introduction

From all the events which occur daily, only a few are deemed to be newsworthy enough to be reported as news in traditional print newspapers or online (Harcup & O'Neill, 2001). Those stories which are picked up by the press usually have special attributes, such as their unexpectedness, their major negative consequences, their effect on the social elite, violent attacks as their topic, eye-catching pictures or their impact on many people (Chermak & Gruenewald, 2006; Galtung & Ruge, 1965).

By analysing the content published by a particular set of mass media outlets (Berelson, 1952) —mass media outlets are those providing written, broadcast or spoken communications that reach a large audience—, the media coverage given to a specific event or topic can be assessed. For example, the influence that news has on the stock market has been investigated (Curme, Zhuo, Moat, & Preis, 2017; Zhang, Zhang, Shen, & Zhang, 2018). Published content can also be used to detect any bias that exists between what is published and reality (Ditton, Chadee, Farrall, Gilchrist, & Bannister, 2004). For instance, it is known that there is more coverage of crimes involving violence or indecency than other crimes (Ditton & Duffy, 1983), with tabloids tending to publish more sensationalist news items (Dickinson, 1993), meaning that what is published in the news differs from reality, in some sense. Also, it has been shown that fake news tends to spread faster than real news since it is more novel, perhaps inspired by shock or fear (Vosoughi, Roy, & Aral, 2018).

The audience is highly selective in their media choices (Lane & Meeker, 2003), their attention span is very limited (Simon, 1971), and collectively, they decide what they want to consume (Morley, 2003). Therefore, although editors of traditional media set a general agenda and coordinate with journalists to decide what is newsworthy, they also consider feedback from their audience (predominantly feedback gleaned after publication, perhaps by years of their audiences' cumulative attention or indifference to certain types of news). Therefore, although this is an area of debate, it might be considered that the audience itself ‘manages the news’ by maintaining or losing interest in a given subject (Downs, 1972). Therefore, any significant discrepancies between reality and what is portrayed by the media reveal the interests of their audience.

1.1. Spatial coverage of the media

The physical distance between individuals and a specific event, or in a more generalised way, any cultural or social ‘distance’, reduces how

* Corresponding author.
E-mail address: rafael.prietocuriel@maths.ox.ac.uk (R. Prieto Curiel).

https://doi.org/10.1016/j.compenvurbsys.2019.02.004
Received 16 August 2018; Received in revised form 13 February 2019; Accepted 13 February 2019
Available online 07 March 2019

connected they feel about it (Tobler, 1970) and so, it also reduces their interest and determines the focus of media on the event. Media data can thus help to uncover the strength of the relations between different regions (Yuan, Liu, & Wei, 2017). It is therefore expected that most of the coverage given by the media is mainly focused on activities and events which are closer to where their audience happens to be. But, what is the balance of coverage of the media with respect to events which are at a similar distance? Are certain locations more newsworthy?

It has been found that people in larger cities tend to suffer more crime (Alves, Ribeiro, & Mendes, 2013; Oliveira, Bastos-Filho, & Menezes, 2017), tend to have a higher income and apply for more patents per capita (Bettencourt, Lobo, Helbing, Kühnert, & West, 2007) and it has also been speculated that they produce less CO2 emissions per person (Fragkias, Lobo, Strumsky, & Seto, 2013). Some socio-economic indicators increase (or decrease) with population size (Bettencourt, Lobo, Strumsky, & West, 2010), for instance, people in larger cities tend to have more social contacts (Schläpfer et al., 2014), they usually walk faster (Bornstein & Bornstein, 1976), and are less likely to migrate (Prieto Curiel, Pappalardo, Gabrielli, & Bishop, 2018). The shared infrastructure of large cities allows them to have fewer petrol stations and big cities have less road surface per person (Bettencourt et al., 2007). Thus, infrastructure, as well as social and economic aspects, ‘scale’ with city size and so, with respect to the media coverage, this begs the question: are large cities more newsworthy than small cities?

The analysis to compare spatial aspects of the media coverage given to various events is challenging since it often reveals differences between the events themselves rather than the social behaviour viewed through the lens of the media. For instance, although they were roughly similar events, it is not straightforward to compare the Bataclan massacre in Paris in November 2015 and the Nice attack in July 2016 for a variety of reasons: the first occurred on a Friday night and the second, on a Thursday, which also happened to be a national holiday (Bastille Day); the first had more fatalities but fewer injured people; and also, on the day of the attack in Nice, Donald Trump announced that he would campaign with Mike Pence for the 2016 US Presidential election and two days later, the Republican National Convention took place, potentially creating a shadow in all but local media with regards to the Nice attack. It is thus hard to make a fair comparison of the media coverage given to the two events.

Rather than considering two or more events, the idea here is to investigate the behaviour of the major news providers following an event that affected a large area and several major conurbations. The attention given by a subset of the most relevant mass media outlets in terms of the size of their audiences which have a national coverage is analysed following an emergency which affected many cities simultaneously, thus, allowing the impact of the population size on media coverage to be detected.

1.2. Temporal coverage of the media

In terms of the temporal aspects, the public attention given to various cultural products follows a consistent pattern over time (Candia, Jara-Figueroa, Rodriguez-Sickert, Barabási, & Hidalgo, 2018; Coman, 2018). In particular, the attention that the audience places on any specific event covered by the media can be categorised into four stages: the pre-problem stage, a discovery stage, a stage of gradual decline in attention and a post-problem stage (Downs, 1972). This attention cycle is closely related to the coverage that the media gives to news related to climate change (McDonald, 2009); the rise and fall of the anti-nuclear movement (Joppke, 1991); terrorism and travel safety (Hall, 2002); and organisational changes in the structures of government (Peters & Hogwood, 1985) among other examples.

Previous studies have attempted to model the evolution of the interest in different published stories on online platforms (Wu & Huberman, 2007), or on social media (Weng, Flammini, Vespignani, & Menczer, 2012; Yang & Leskovec, 2011), focusing only on the attention given by the readers (or the users) measured by the ‘upvotes’, ‘shares of memes’ or ‘hashtags’, rather than by the news providers. Here, we consider how long an event remains in the media spotlight.

The analysis of the temporal aspects of the coverage given by the media is also challenging since it often reveals the evolution of the coverage of topics rather than specific events, so it is usually hard to identify a starting point of interest in a particular topic. For instance, the public attention to climate change tends to display cyclic variations (McDonald, 2009), so there is no clear starting point for climate-related media. Also, the number of stories related to terrorism had a spike after the attacks of Madrid and London (Petersen, 2009) but this topic is never far from the headlines, so there is no clear starting point for terrorist-related media either.

Despite having clearer temporal bounds, the coverage from the media of specific events (and not of a more general topic) is relatively small and tends to vanish after a few days and so it is not usually possible to detect any structural decay of the amount of coverage.

Rather than considering a specific topic or the audience's reaction to a published story, the idea here is to consider the evolution of the coverage from the mass media following a specific event of great importance that captured the audience's interest for a relatively long period, and the evolution of the spatial distribution of the media coverage in the weeks after this event.

1.3. Quantifying media coverage

Detecting the content that a media outlet publishes is not straightforward and, because of the new ways of delivering media, this now requires new ways to be established to quantify how the attention is distributed to different issues.

1.3.1. Approaches to quantifying media coverage

A common way to quantify and to analyse the focus of media is to measure the area in printed newspapers and magazines (or the minutes on radio or TV) devoted to a particular story. Using this approach, for instance, it was discovered that newspapers concentrate on crimes which have a sexual or a violent component (Chadee & Ditton, 2005; Ditton & Duffy, 1983) highlighted by the fact that in the USA homicide (murder) makes up only 0.02% of all crimes suffered but these form 30% of crime stories in the newspapers (Liska & Baccaglini, 1990). Similar studies have shown that climate change coverage differs strongly between countries, with a higher attention given in carbon dependent countries with commitments under the Kyoto Protocol (Schmidt, Ivanova, & Schäfer, 2013). Also, by using this same approach, a decay in the coverage of terrorism news given by a newspaper between 2000 and 2007 (so, the period after the S11 attack) was observed (Petersen, 2009).

1.3.2. Moving towards online media

Over recent years, many traditional media outlets such as newspapers or magazines have dramatically changed their methods of delivery. For example, now, less than 17% of The Guardian's audience actually reads the printed edition of that newspaper (PressGazette, n.d.). The main media channels with an informative focus, have moved to online streams and hence, the ways to analyse the focus of the media have also changed.

Therefore, one possible strategy would be to count the number of different URLs published by a newspaper, for example. To help with this analysis, there are resources for searching links published by different online media outlets, such as the GDELT Project (Leetaru & Schrodt, 2013), so that outputs from different sites can be analysed to reveal items related to different topics, for example, to climate change (Olteanu, Castillo, Diakopoulos, & Aberer, 2015). However, during major events, different URLs are not necessarily associated to different news items (for instance, an entry corresponding to one particular URL can be updated several times as the events unfold), thus it becomes difficult to count them and so it becomes hard to measure the media coverage based solely on the number of URLs. Furthermore, identifying all the URLs published by a website can also be challenging, as they are usually non-trivially divided into sections, subsections, columns and more.

1.3.3. Using social media to measure the coverage of the most popular online newspapers

The collection and classification of tweets to detect a general trend has been carried out before (Amato et al., 2017; Kounadi, Lampoltshammer, Groff, Sitko, & Leitner, 2015; Pak & Paroubek, 2010) and this has enabled the detection of flows of national politics and news (Ausserhofer & Maireder, 2013), crime hotspots (Malleson & Andresen, 2015), exposure to cross-ideological content (Himelboim, McCreery, & Smith, 2013), activism (Xu, Sang, Blasiola, & Park, 2014) and more. Understanding the dynamics of social networks has been proven to be crucial for the management of crises such as environmental natural hazards or diseases (Cvetojevic & Hochmair, 2018).

In the current work, the coverage given by the most popular national online newspapers, in terms of the size of the audience, is considered. It is assumed that each time an update to the corresponding newspaper's website is issued, this is usually announced, at least once, via a related post on its associated Twitter account. The reasoning supporting this assumption, is that with little effort or cost (compared to that of producing the news update), mass media outlets can reach a broader audience by promoting their updates on their social media accounts. Thus, the online newspapers have an incentive to promote their contents on Twitter; contents that would be otherwise more difficult for the audience to reach. It is also assumed here that by taking into account the online newspapers with largest audiences, a significant proportion of the national written news output is being considered. Moreover, these media outlets have different interests (commercial and public) and target different audiences. Therefore, their output can be regarded as representative of the overall written news activity.

Collecting Twitter data is relatively easy and less time-consuming than other methods to determine media coverage. Thus, although not a perfect metric, the number of tweets published on a newspaper's Twitter account about a particular subject is used as a proxy to detect the media coverage. If a specific news item is posted on Twitter several times by one account, then this indicates what the outlet considers this to be important and also, news items which are not posted on Twitter are likely to be less relevant for the newspaper.

1.4. The S19 earthquake in Mexico, 2017

In order to better understand the spatial and temporal diffusion of the amount of media coverage in the aftermath of a major event, news items published only by Mexican mass media outlets which have a national coverage were investigated. The event was an earthquake which hit parts of Mexico on the 19th of September of 2017 (hereafter referred to as S19). The earthquake was mainly felt in the central part of Mexico simultaneously affecting cities of various sizes: a large city (Mexico City, population of 22 million), medium-size cities (such as Puebla and Toluca, population above 2 million), small towns (such as Jojutla, population just over 40,000) and rural areas in the state of Puebla and Morelos.

The magnitude of the earthquake (7.1 Mw) caused significant damage which ensured that it remained in the media spotlight for several weeks. Although this is, perhaps, a special event, it does allow us to consider how the coverage of the major national outlets across several cities changed as a result of the earthquake since it affected many cities concurrently. This said, the impact of the disaster was not homogeneously distributed over the country. In Mexico City, there were 228 deaths and in the state of Morelos (population of 1.8 million), there were 75 so that, taking into account the population size of both, the earthquake was four times more deadly in Morelos than in Mexico City. The questions we ask though are: how is the coverage given by the most popular Mexican online newspapers, per person, distributed among different cities whose size of population varies? and, how does this coverage decay over time?

2. Data and methods

2.1. Identifying media channels

There are a large number of mass media outlets in Mexico each offering an informative approach to news, but only a few are taken into account in this study. For the analysis, only the most popular —by size of the audience— national online newspapers are considered, since they are assumed to be representative of the overall written news activity produced by the national mass media.

Given the convenience of retrieving data from Twitter, the coverage given by the most popular national online newspapers in Mexico is measured through their Twitter activity, which is assumed to be a proxy of the newspapers' coverage. It is also assumed that the number of followers of the associated Twitter accounts can be used to estimate the size of the newspapers' audience. Tweets that can be associated to large media companies have been shown to receive more attention per follower than tweets posted by other users, suggesting that the former have already built up their follower network and trustworthiness (Cvetojevic & Hochmair, 2018).

The Twitter accounts belonging to the 19 most popular national newspapers in Mexico in their online version are considered. For more information on the choice of newspapers, see the Supplementary material. These are then sorted in decreasing order of popularity according to their number of followers on Twitter. For account A_i, with $i = 1, 2, \dots, 19$, the number of Twitter followers is denoted as F_i. The most popular account has 7.8 million followers (so $F_1 = 7.8$m) whilst the least popular account has 50 thousand followers (so $F_{19} = 50{,}000$). In total, the 19 accounts have 41.3 million followers, although it should be noted that a Twitter user might follow more than one newspaper, there being only around 9 million Twitter users in Mexico (whose population is close to 130 million).

2.2. Collecting tweets from selected accounts

In order to collect the tweets, the Twitter API "Get Tweet timelines" (Twitter, n.d.) was used. This tool allows the tweets from the specified accounts to be retrieved. The retrieval of tweets was carried out periodically and thus the full set of tweets from each account considered here was collected.

To establish a baseline, 18,083 tweets were collected for a full week before the S19 earthquake. It is important to add that an earthquake of smaller magnitude occurred on the 7th of September. This prompted the selection of the week between the 29th August and the 4th September 2017 instead of the week immediately prior to the S19. Then, 73,389 tweets were collected between the 19th September and the 17th October 2017, corresponding to the four weeks starting from the day of the S19 earthquake. In total, 91,472 tweets were collected from the 19 accounts which include all the tweets posted by those accounts over the five weeks in which the data was collected.

2.3. Cities in the media

For the purposes of this analysis, a group of 30 cities were considered. In total, 15 of these cities were affected by the earthquake. The attention received by the 15 affected cities was measured during the four weeks after the S19 earthquake. These cities were selected based on two different criteria: their population had to be more than 20,000 and the local measure of intensity for the S19 earthquake had to be moderate or higher using the Modified Mercalli Intensity scale.

Additionally, 15 more cities that were not affected by the earthquake were considered for the analysis during the baseline week before the earthquake. These were selected to have a population of at least 450,000 inhabitants, as they were big enough to be featured with a relatively high frequency in national news. Additional information on how cities for the study were selected is available in the Supplementary material.

For each city, its population size is denoted by P_j, where now $j = 1, 2, \ldots, 30$, and these are ordered by size so that $P_1 = 22$ million inhabitants (the metropolitan area of Mexico City) and $P_{30} = 30$ thousand inhabitants (Tlaquiltenango, in the state of Morelos).

2.4. Classifying a tweet as being related to a specific city or to the earthquake

Tweets can be examined to see if they contain specific words (Dodds, Harris, Kloumann, Bliss, & Danforth, 2011) and so here the collected tweets were classified as being related to the earthquake, a city, or both, depending on whether or not they contained words or hashtags associated with the earthquake or a specific city. To do this, two lists of terms were created. One contained terms related to the earthquake such as *terremoto* (earthquake), daños (damages), *#FuerzaMexico* or *damnificado* (affected). The other list, contained terms related to the cities, such as the name of the city itself, the names of the city mayors and the names of significant places or buildings —this last aspect is particularly important in the classification of news items related to Mexico City, as tweets often refer to a specific location within the city, but without mentioning the city by name. Our lists can be found at https://doi.org/10.6084/m9.figshare.7347320.v1.

As these lists of predefined terms may not be exhaustive, the tweets that did not match any of the terms were manually examined and classified a posteriori. This ensures no false negatives in the classification process. The false positives were not controlled, but their number may be reduced due to the context in which the tweets were generated: it is assumed that a tweet matching 'earthquake' or 'reconstruction' during the week after an earthquake is related to the natural disaster, particularly since only tweets belonging to online newspapers are analysed.

Another point to take into account is that none of the 19 national media outlets that were used for the analysis share their location, therefore tweet classification is purely made by checking if the tweets contain any city-related terms. If instead, we had considered individual users who felt the earthquake, their location would have been relevant. For example, elsewhere in order to classify the online media users according to their behaviour on social networks, their IP addresses have been used to identify their location (Zelenkauskaite & Balduccini, 2017).

From the 73,389 tweets posted during the four-week period after the earthquake, more than one-third of these contained information related to the earthquake.

2.5. Defining the media coverage index

To quantify the amount of media coverage received by each of the 30 cities from the Twitter accounts belonging to the 19 national newspapers in Mexico with the highest popularity in their online form, a media coverage index is defined.

The number of tweets published by each of these accounts over a certain period of time, denoted as T_i with $i = 1, \ldots, 19$, can be established. From these posts, the number which includes a reference to city j can also be identified and denoted as $N_{i,j}$, so that, for instance, $N_{2,7}$ represents the number of tweets published by account A_2 which contain a reference to city 7. In addition to quantifying the media coverage given to a certain city, we might be interested in quantifying the media coverage given to a city because of a certain topic, like the earthquake in this case. Then $N_{i,j}$ must be taken as the number of tweets published by account i which contain a reference to both the city j and the topic in question.

Some accounts publish more frequently than others with only slight changes in the content, often for marketing reasons. In order to counter this bias, the quotient of $N_{i,j}/T_i$ is formed. To quantify the coverage given by account A_i to city j in relation to the size of its influence, the media coverage index I_j for city j is defined by multiplying this quotient by the number of its followers F_i and summing over all the accounts

$$I_j = \frac{1}{\overline{F}^{(19)}} \sum_{i=1}^{19} F_i \frac{N_{i,j}}{T_i} \tag{1}$$

where $\overline{F}^{(n)} = \sum_{i=1}^{n} F_i$ and therefore, the factor $1/\overline{F}^{(n)}$ simply rescales the index so it is dimensionless and lies between 0 and 1. Although the current work focuses its attention on the main Mexican newspapers' Twitter accounts, the definition of the media coverage index is general and could be used for other countries and other types of media. Considering more (or fewer) accounts would not change the index significantly since the media coverage index is defined in such a way that it is robust to changes in the number of accounts when n is large enough (see the Supplementary material).

3. Results

3.1. Number of tweets devoted to the earthquake

On a normal midweek day, prior to the earthquake, the 19 media accounts considered here publish 2,813 ± 229 tweets, decreasing to 2,008 ± 249 over the weekends. There are also fluctuations within a 24 hours period with fewer tweets released at night. On the day

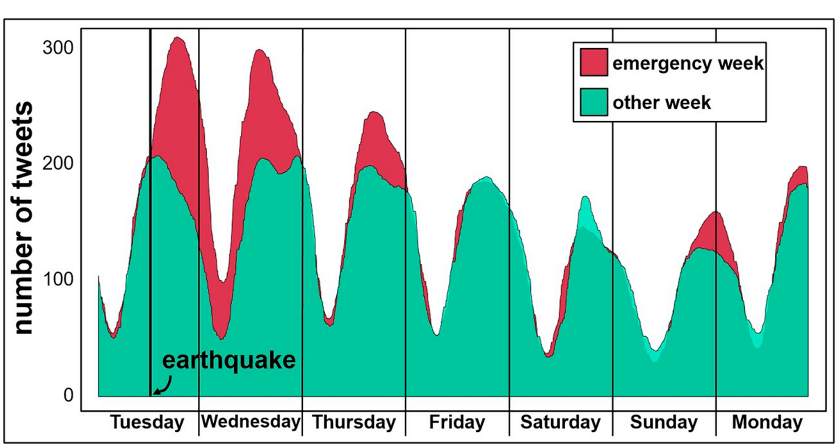

Fig. 1. Evolution of the number of tweets published by the 19 selected media outlets smoothed over one-hour long time windows, comparing a regular week and the earthquake week (starting from Tuesday, the day of the earthquake). The number of tweets has daily fluctuations, with a decrease during the weekend and during the night. During the week of the earthquake, the media published more posts but only during the first 72 hours after the earthquake. Thereafter, the number of tweets returned to its regular daily fluctuations.

immediately after the earthquake, the number of tweets increased to 3,316, and to 4,415 the following day (an increase of 17% and 56% compared to a regular working day). However, the publication rate returned to its usual value within three or four days (Fig. 1), which overlays a typical normal week in blue and the week immediately after the earthquake in red.

Tweets can be published at any time, and so it is common to find high concentrations of tweets at certain times of the day but also periods when no posts are published at all. For the analysis, a technique inspired by moving average methods is used to smooth the data over time (see Supplementary material) and so, hereafter, smoothed variables will be considered.

Let $\nu(t)$ be the number of tweets published by the media at time t, measured in days. Then, if $t = 0$ is set to be the moment the earthquake struck, $t = 1$ would be the day after the earthquake. Let $\eta(t)$ be the number of tweets which are related to the earthquake, so that $\eta(0)$ is the number of tweets related to the earthquake immediately after it hit, and $\eta(t) = 0$ for $t < 0$. Both the number of tweets and the number of earthquake-related tweets have some daily and weekly fluctuations. Let $a(t) = \eta(t)/\nu(t)$ be the proportion of tweets devoted to the earthquake at time t, with $a(t) = 0$ for $t < 0$. Then, $a(t)$ can be interpreted as the 'spotlight' or equivalently, the probability that a random tweet published around time t was related to the earthquake.

Almost immediately after the earthquake, nearly 83% of the tweets contained information about the earthquake so that although the peak in earthquake-related tweets did not occur strictly at $t = 0$, $a(0)$ is set to $a(0) = a_0 = 0.83$ for simplicity. During the following hours, the percentage of the attention of the media placed on the earthquake remained similarly high, but thereafter, the interest diminished with an exponential decay (Fig. 2).

To model the decay in media interest, an assumption can be made such that for $t > 0$ it decreases exponentially:

$$a(t) = a_0 \psi^t \tag{2}$$

with a factor $a_0 = 0.83$, as this is the initial proportion of news devoted to the earthquake. Note that, the exponential decay function has been expressed with the parameter ψ as base, which needs to be estimated from the data. With the time t measured in days, ψ can be interpreted as the proportion of news related to the earthquake with respect to the previous day. The estimated value for ψ is $\widehat{\psi} = 0.9196$ (obtained by nonlinear minimisation of the mean squared error). This value $\widehat{\psi}$ indicates that each day, there is approximately 8% less earthquake-related coverage than the previous day. Equivalently, after eight days, the coverage of the media drops to half with respect to the initial day ($0.5 \approx \widehat{\psi}^8 = 0.9196^8$).

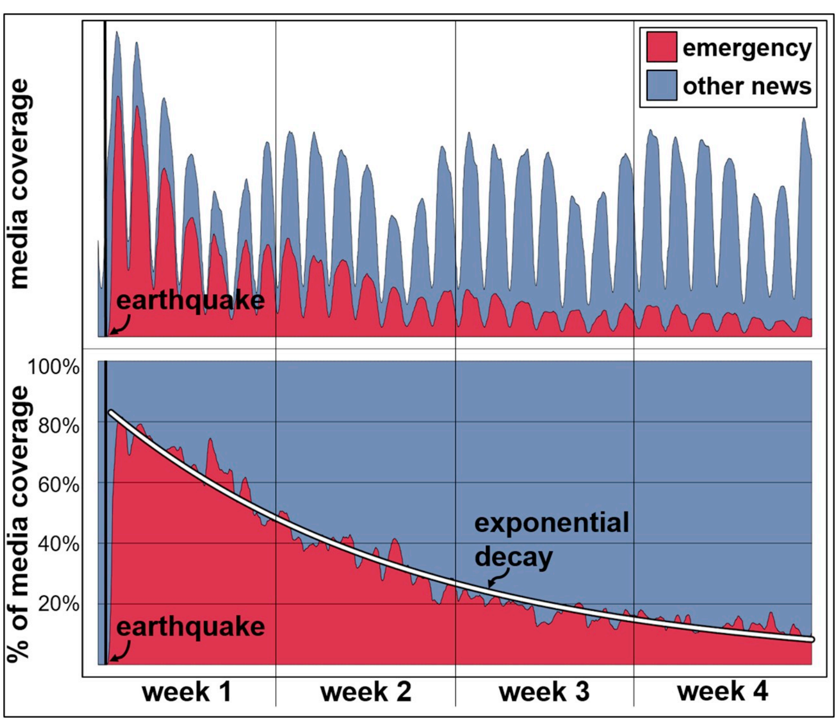

Fig. 2. Changing media coverage. The upper panel shows the evolution of the number of tweets published by the 19 media outlets $\nu(t)$ smoothed over one-hour long time windows, as well as the number of earthquake-related tweets, $\eta(t)$, also smoothed over time between 19th September to the 19th October 2017. The lower panel shows the portion of tweets devoted to the earthquake $a(t)$ overlaid with and an exponential decay curve.

3.2. Location of the spotlight: cities being mentioned on the earthquake news

During a regular (pre-earthquake) week, for every tweet that mentions a city that is around 1 million inhabitants (Querétaro, Mérida or Cuernavaca), 70 tweets are found related to Mexico City, which is 22 times the size. Therefore, on a per capita basis, the media coverage index for Mexico City is 3.2 times larger than for a city with only 1 million people. Results show that there is a superlinear behaviour with respect to media coverage. However, just after the earthquake, it was noted that the city of Celaya, with just half million inhabitants, had the same number of tweets per person as Mexico City, so that there might be outliers, perhaps due to some specific situations on that city.

To formally compare the coverage of the media focused on the 30 cities, a power law equation between the media coverage index I_j corresponding to city j, and the population of each city P_j is assumed to be given by

$$I_j = \alpha P_j^{\beta} \tag{3}$$

where parameters $\alpha > 0$ and $\beta > 0$ are to be determined from the data. The estimated coverage per person varies as $P_j^{\beta-1}$ (dividing both sides of Eq. 3 by the population of the city j, P_j) and so, if $\beta > 1$, the coverage increases with city size, while if $\beta < 1$, the coverage decreases with city size. If the data reveals a value of β close to 1, then the media coverage index is not affected by the city size.

Data for the coverage placed on each city during a typical week before the earthquake gives a coefficient of $\widehat{\beta}_{typ} = 1.385 \pm 0.118$, meaning that, during a typical week, there is a superlinear scaling of the coverage from the media to each city according to the size of their populations (Fig. 3). For the data corresponding to the first week after the earthquake, the coefficient is found to be $\widehat{\beta}_{eqk} = 0.7526 \pm 0.6728$,

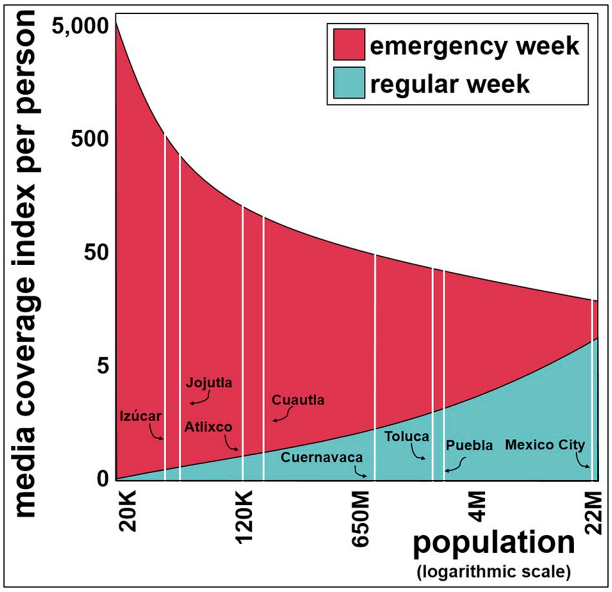

Fig. 3. Media coverage index, per person, for cities of different size. During a typical pre-earthquake week, the media coverage index per person scales superlinearly, so large cities received greater coverage per person. However, during the immediate post-earthquake period, the media coverage index does not scale superlinearly anymore, meaning that national media gave greater coverage per person to smaller cities.

meaning that the strong superlinear behaviour observed during the baseline week (before the earthquake) is no longer observed. This changes back to the clear superlinear scaling as time passes (Fig. 4).

Considering the scaling coefficient of the media coverage now as a continuous function of time $\beta(t)$, again with t measured in days, results show that by one week after the earthquake, $\beta(7) \approx 0.62$ (the lowest coefficient observed after the earthquake) showing that coverage is given to smaller cities but this behaviour changes by the end of the third week when $\beta(21) \approx 1$. Then $\beta(t)$ goes back to nearly the values observed before the earthquake (Fig. 5), so larger cities are again the ones that create more news.

4. Conclusions

It has been found that the producers of the most popular online newspapers with national coverage in Mexico and hence, also their consumers, follow an emergent collective behaviour by which the amount of coverage that they give to an event that affected a range of Mexican cities decays over time. Additionally, the amount of attention given to particular cities depends on the size of the city (as measured by the size of the population), with this dependency being fundamentally different just following the event.

To quantitatively analyse the news spotlight given to a specific subject —in this case, an earthquake in Mexico— and to cities of different size, a media coverage index has been defined by taking into account the amount of coverage given by the 19 most relevant national online newspapers in Mexico, in terms of the size of their audience, since these are assumed to provide a good representation of the overall written news activity from the national media outlets. The amount of coverage given by each outlet is measured by counting the number of tweets related to the subject that are posted via the Twitter accounts corresponding to the newspapers. The assumption is that what appears in the Twitter accounts is a reflection of what also appears in their online, printed or broadcast material but this goodness of fit has not been formally assessed. The number of tweets corresponding to each Twitter account is weighted by the number of followers of the account, which is assumed to be a proxy for the size of the outlet's audience.

4.1. Rapid decay of the media attention

Results show that the media coverage as measured here had an initial burst after the event, but it decayed exponentially over time. Roughly, every day the event has 8% less attention than the previous day. Similar to cultural products, such as music or academic papers (Candia et al., 2018; Coman, 2018), media coverage has a distinctive decay.

Although there might be other specific aspects which determine how much coverage is given to other subjects, a similar behaviour is expected for the interest that media shows in situations other than this earthquake. The audience has an initial interest denoted by A_0, and then, as t days go by, the attention placed on the same event, $A(t)$, decays according to $A(t) = A_0\psi^t$, where ψ depends on factors such as locality or whether sports, stars or celebrities are involved. The earthquake was a particularly significant event and therefore there was a major initial post-earthquake interest but it is likely that a similar decay will exist for other events albeit with a lower initial peak.

4.2. The coverage of the media is not homogeneously distributed

The earthquake in Mexico provided a special event to quantitatively measure national media coverage given by the most popular online newspapers over time but also to consider a spatial factor. The earthquake affected a large area, including a number of cities and yet the media spotlight on different cities varied.

Results show that large cities are more newsworthy, but a major event might change it. In the absence of the earthquake, there is usually a superlinear behaviour in the dependence of the media coverage index with regards to the size of the population of the city. Using a similar approach it has been previously shown that people from large cities produce more new patents ($\beta = 1.27$); tend to suffer more serious crime ($\beta = 1.16$); have a higher wage ($\beta = 1.12$) and consume more electricity ($\beta = 1.07$) (Bettencourt et al., 2007). Before the earthquake, the superlinear scaling coefficient of $\beta = 1.3768$ is surprisingly high

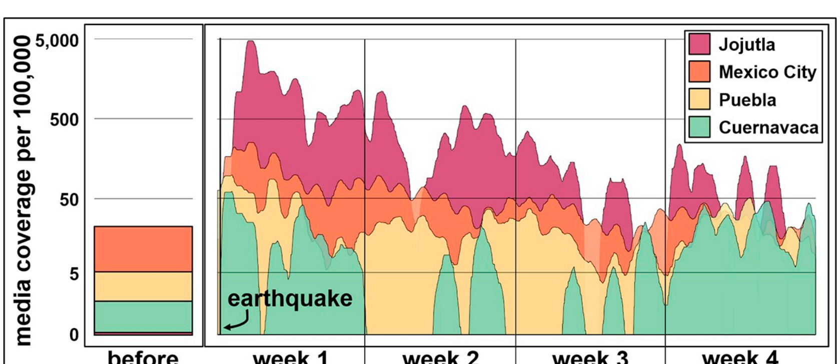

Fig. 4. Evolution of the media coverage as per the number of tweets, per 100,000 persons, for four cities of various sizes. During a typical pre-earthquake week, large cities received a greater number of tweets per person. During the first few weeks after the earthquake, the number of tweets per person referring to smaller cities like Jojutla overtook the number of tweets per person referring to Mexico City. The distribution of interest gets closer to the levels observed during the pre-earthquake week as time goes by.

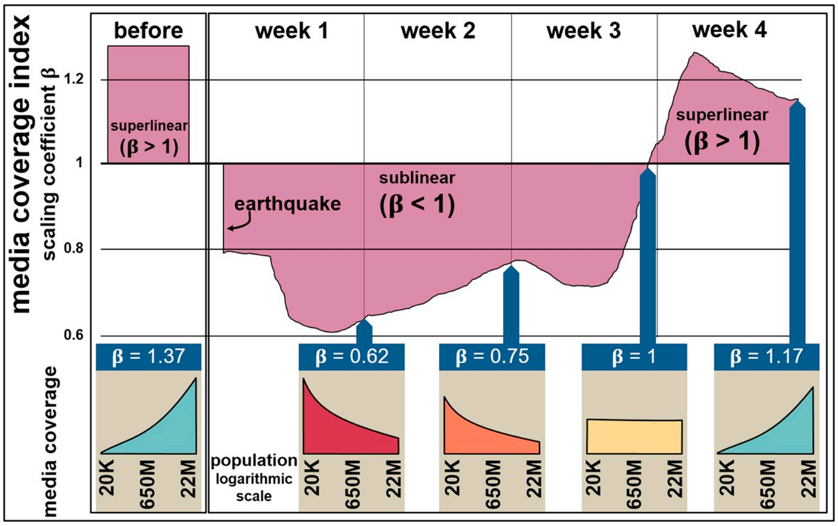

Fig. 5. Evolution of the scaling coefficient β. During a typical week, large cities receive higher media attention per person, therefore, the media coverage index scaling coefficient β is greater than 1, but during the period after the earthquake, the media coverage index per person does not scale superlinearly with city size anymore, so $\beta < 1$. A few days after the earthquake, the scaling of the media coverage index comes back to normality.

meaning that events from large cities are seemingly more newsworthy than events in smaller cities. Immediately after the earthquake, the scaling is no longer superlinear and so small cities such as Jojutla, which previously were rarely mentioned, received considerably more attention. However, this change to the spatial distribution of interest is a short-term effect since, as the general interest in the earthquake decays, the spatial distribution of the media coverage also comes back to its usual form, i.e. four weeks after the earthquake large cities receive again more attention per capita from the media.

A more refined version could be achieved by incorporating a measure of the impact of the earthquake within the definition of the media coverage index, but this is left for future work.

It is interesting to note that in some cases, world capitals can be considered as outliers (popularly known as dragon-kings (Sornette, 2009)) and might affect the whole urban system when it comes to the application of scaling analysis (as observed in the UK, for instance (Arcaute et al., 2015)). However, in this case, the results after the earthquake were compared against the baseline week before the earthquake, so it is possible to see how the attention of the media was distributed across cities before the event. The results show that the scaling of attention changed rapidly.

The trends detected here, based on the analysis of the Twitter accounts belonging to the newspapers which provide national coverage and with the highest popularity in their online platforms, are likely to be reflected also in traditional printed media and other news outlets, including those with a regional or global outlook. The methodology suggested here can be applied to other regions of the world and to other equally suitable events.

Acknowledgements

RPC acknowledges the support of the government of Mexico via a Conacyt Scholar. CCA acknowledges the support of University College London. MTP acknowledges the support of the government of Mexico via a Conacyt Scholar. This article was completed with support from the PEAK Urban programme, funded by UKRI's Global Challenge Research Fund, Grant Ref: ES/P011055/1.

Declaration of interest

The authors declare that they have no competing interests.

Appendix A. Supplementary material

A.1. Selection of media outlets

It is important to note that although all the selected media channels have a similar approach, there is still variability among them: some have a more commercial purpose and others focus more on informing citizens. For example, *Televisa* is part of a large consortium which also has many TV channels, and *OnceTV* is broadcast by a University with subsidised resources. However, the analysis in the following sections shows that the same trends are observed in all the outlets, with similar patterns of decay and scaling coefficients. Thus, the results are general, rather than specific for only one particular type of outlet.

A.2. Selection of cities

Firstly, 15 cities affected by the S19 earthquake were selected for analysis, based on the intensity of the earthquake and the size of their population. According to the US Geological Service Report, the S19 earthquake (USGS Event ID: us2000ar20) had a magnitude of M 7.1 and was centred 1km ESE of Ayutla. For the current study, we selected cities that suffered the biggest impact from the earthquake, since these would also be the ones likely to generate more news items. Although it might seem that these cities must be those that are closest to the epicentre, this is not always true, since the impact also depends on other factors, such as the local geological conditions as well as the quality of the local infrastructure and the types of structures within the city (high rise, old style buildings, etc.).

#	Outlet	Twitter account	Type	Followers (thousands)
1	Aristegui Noticias	@AristeguiOnline	Commercial	7,742
2	El Universal	@El Universal Mx	Commercial	4,668
3	Proceso	@revistaproceso	Commercial	4,595
4	Milenio	@Milenio	Commercial	4,018
5	Sopitas	@sopitas	Commercial	2,773
6	Televisa	@NTelevisa com	Commercial	2,710
7	Reforma	@Reforma	Commercial	2,460
8	La Jornada	@lajornadaonline	Commercial	2,022
9	Animal Pol´ıtico	@Pajaropolitico	Commercial	1,722
10	Excelsior	@Excelsior	Commercial	1,505
11	Sin Embargo	@SinEmbargoMX	Commercial	1,145
12	Pictoline	@pictoline	Commercial	1,097
13	El Financiero	@ElFinanciero Mx	Commercial	1,016
14	Publimetro	@PublimetroMX	Commercial	689
15	El Economista	@eleconomista	Commercial	350
16	UnoTV	@UnoNoticias	Commercial	321
17	SDP	@sdpnoticias	Commercial	319
18	La Silla Rota	@la sillarota	Commercial	287
19	Once Noticias	@OnceNoticiasTV	Public	68

Thus, only cities where the modified Mercalli Intensity scale was moderate or higher were considered. In earthquake dynamics this scale quantifies the local mea- sure of intensity according to the responses from people and buildings rather than the raw excitation force, so it measures the response rather than the force.

In order to analyse the scaling of media attention during an ordinary week, 15 cities which did not feel the earthquake, but which have a population of at least 450,000 inhabitants, were also considered. Smaller non-affected cities were not con- sidered since the number of news items related to these is not significant. Cities located on the border between Mexico and the United States, despite having a pop- ulation larger than 450,000 inhabitants, were not considered as part of the study either since they might receive media attention from outlets that use English as their working language. This is the case of Tijuana, Mexicali, Ciudad Ju´arez, Reynosa, Matamoros and Nuevo Laredo. Thus, in total, 30 cities were considered for the study. It is important to note that even though Twitter penetration can be higher in bigger cities, the results presented here do not depend on this observation. Assuming that the Twitter content is similar to the content in other platforms belonging to the same newspaper, if media channels decide to publish more news items about the big cities due to the higher number of followers there, then this is considered to be part of the scaling effects.

A database which contains the cities considered for the current study, and the population within their metropolitan area, can be found at https://doi.org/10.6084/m9.figshare.7347320.v1

A.3. Classification of tweets

Detecting whether a particular tweet is related to the earthquake or to a specific city poses a challenge, as it is not possible to manually read each post. To facilitate the classification of the tweets they were pre-processed to remove punctuation marks, prepositions and articles, which do not provide information on whether a tweet is related to a city or to the earthquake.

The tweets were then classified as being related to the earthquake and/or to one or more of the Mexican cities under consideration by checking if the tweets contained any of the terms contained in the two lists of words: one containing earthquake- related terms, and the other containing city-related terms.

Once an initial classification by successive look-ups of the terms included in the lists was carried out, it was relatively quick to manually check, tweet by tweet, that the classification was correct. When it comes to both selecting earthquake or city-related tweets the method proved to be very accurate, with very few amend- ments needed. For a bigger database, it would not be possible to do this and more advanced techniques would be required. Particularly, the inclusion of names of collapsed buildings and specific events made the selection of words and tweets easier. For instance, after a building, 'AO286' collapsed, tweets containing that word were also identified as earthquake-related.

It is fully acknowledged that it is possible to apply more advanced machine learning methods to perform this classification but, given that the earthquake topic is easily detected in tweets and the amount of tweets is relatively small so they can be manually checked to assess the accuracy of the classification, these are not needed for this one-off study and the miss-classification errors were kept to a minimum.

A database which contains the list of words which were used to detect whether a tweet is related to a city and related to the earthquake can be found at https://doi.org/10.6084/m9.figshare.7347320.v1

A.4. Smoothing the tweets over "continuous" time

Since tweets might be posted at any time, the count of the number of tweets posted up to a certain time needs to be smoothed over time for further analysis. Consider, as a starting point, a partition that divides the time interval under consideration starting from a selected initial time, into periods denoted by m_i^0, with $i = 1, 2, 3, \ldots$, each with a width w_i^0. Both the count of the total number of tweets and the number of tweets related to the earthquake that were posted during each period, N_i and E_i respectively, can then be established. In order to produce a smooth version of the evolution of the number of tweets (total and earthquake-related), a process inspired by moving average methods is applied. A number p of additional non-overlapping partitions of the total time in- terval are generated each with different period widths and different random starting positions. The counting of tweets is performed again for each new partition. Segments of the partition without news are ignored as no information with respect to the attention of the media on these segments is provided.

A new value is assigned to each of the original time periods m_i^0. This value is obtained by considering the average of the total number of tweets corresponding to the periods that, in each of the additional partitions, overlap totally or partially with the original periods. By representing all these values against time, it is possible to obtain smooth-looking versions of the total number of published tweets $\nu(t)$, the number of earthquake-related tweets $\eta(t)$ and the proportion of earthquake tweets$a(t)$. Although these magnitudes only take discrete values for each of the original time periods, if each w_i^0 is small enough, the evolution of $\nu(t)$, $\eta(t)$ and $a(t)$ will be close to continuous.

A.5. Robustness of the media attention index

One question that arises from the definition of the media attention index is if it is robust under a change in the number of Twitter accounts. In order to study this, changes to the media attention index are investigated by considering, say $n-1$ instead of n Twitter accounts. Let the media attention index for the n accounts focused on city j be $I_j(n)$, where now the number of accounts is made explicit in the notation. This can be rewritten in terms of $I_j(n-1)$:

$$\mathrm{Ij(n)} = \frac{1}{\overline{F}^{(n)}}\sum_{i=1}^{n} F_i \frac{N_{i,j}}{T_i} = \frac{\overline{F}^{(n-1)}}{\overline{F}^{(n)}} I_j(n-1) + \frac{F_n}{\overline{F}^{(n)}}\frac{N_{n,j}}{T_n} \tag{A.1}$$

One important thing to note is that, since $\overline{F}^{(n)} = \overline{F}^{(n-1)} + F_n$ and $\overline{F}^{(n-1)} >> F_n$ when n is large enough, the quantity $\frac{\overline{F}^{(n-1)}}{\overline{F}^{(n)}}$ in Eq. (A.1) is close to 1, and $\frac{F_n}{F_{(n)}}$ is close to 0.

It is also important to realise that in the extreme case when this additional account A_n does not post any tweets related to city j, then $I_j(n) = \frac{\overline{F}^{(n-1)}}{\overline{F}^{(n)}} I_j(n-1)$. Therefore, since all the terms are positive the value of $I_j(n)$ is bounded from below by $\frac{\overline{F}^{(n-1)}}{\overline{F}^{(n)}} I_j(n-1)$.

It remains to be seen which one from all the possible upper bounds, would be the largest, so that the whole range of values for $I_j(n)$ are taken into account. If we imagine that F_i and T_i are fixed for i = 1,..., n, the largest upper bound corresponds to the situation where all the tweets posted by A_n refer to city j, hence $\frac{N_{n,j}}{T_n} = 1$. Thus,

$$\frac{\overline{F}^{(n-1)}}{\overline{F}^{(n)}} I_j(n-1) \le I_j(n-1) \le \frac{\overline{F}^{(n-1)}}{\overline{F}^{(n)}} I_j(n-1) + \frac{F_n}{\overline{F}^{(n)}} \tag{A.2}$$

For the media attention index to be a robust quantity with respect to the number of Twitter accounts, the $I_j(n-1)$ must be shown not differ significantly from $I_j(n)$. By looking at the lower bound in Eq. (A.2) this is seen to be the case since $\frac{\overline{F}^{(n-1)}}{\overline{F}^{(n)}} I_j(n-1) \approx I_j(n-1)$ the upper bound, there is one realistic assumption that can be made: if A_n places all its attention on city j, then it is likely that all the other accounts *A1*,..., *An-1* also give most of their attention to city j. Therefore, the term $\frac{F_n}{\overline{F}^{(n)}}$ will be small compared to $\frac{\overline{F}^{(n-1)}}{\overline{F}^{(n)}} I_j(n-1) \approx I_j(n-1)$, as $I_j(n-1)$ will be close to 1. Following this reasoning, it is possible to see that the more accounts considered (so the bigger n is), then the effect that A_n has on $I_j(n)$ is lessened and the closer $I_j(n)$ is to $I_j(n-1)$.

To illustrate the robust nature of the index, the media attention index corresponding to Mexico City during the first week after the earthquake can be obtained from the top 19 media outlets (or from the top 18, leaving one out). In the first case, the result is 0.594436 while if the least popular outlet is removed, then the index has a value of 0.594441, which is a change of less than 0.001%. Even if the most popular from the 19 media outlets were to be removed, the media attention index would be 0.628648, and so the change would be 5.442%.

A.6. Table of coeffcients

Table 1
Coeffcients of the memory decay model.

	Lower bound	Estimate	Upper bound
Initial interest of the media, *a0*	0.81061	0.8300	0.83519
Memory decay of the media, ψ	0.91785	0.9196	0.92118

Table 2
Coeffcients of scaling of media coverage during a regular week

	Estimate	Error
Regular week β	1.3855	0.1181
Earthquake period β	0.7526	0.6728

References

Alves, L. G., Ribeiro, H. V., & Mendes, R. S. (2013). Scaling laws in the dynamics of crime growth rate. *Physica A: Statistical Mechanics and its Applications, 392*(11), 2672–2679.

Amato, G., Bolettieri, P., Monteiro de Lira, V., Muntean, C. I., Perego, R., & Renso, C. (2017). Social media image recognition for food trend analysis. *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval* (pp. 1333–1336). ACM. https://doi.org/10.1145/3077136.3084142.

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., & Batty, M. (2015). Constructing cities, deconstructing scaling laws. *Journal of the Royal Society Interface, 12*(102), https://doi.org/10.1098/rsif.2014.0745http://rsif.royalsocietypublishing.org/content/12/102/20140745.full.pdf.

Ausserhofer, J., & Maireder, A. (2013). National politics on Twitter: Structures and topics of a networked public sphere. *Information, Communication & Society, 16*(3), 291–314. https://doi.org/10.1080/1369118X.2012.756050.

Berelson, B. (1952). Content analysis in communication research. *The Annals of the American Academy of Political and Social Science*. https://doi.org/10.1086/617924.

Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences, 104*(17), 7301–7306. https://doi.org/10.1073/pnas.0610172104.

Bettencourt, L. M., Lobo, J., Strumsky, D., & West, G. B. (2010). Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLoS ONE, 5*(11), 13541. https://doi.org/10.1371/journal.pone.0013541.

Bornstein, M. H., & Bornstein, H. G. (1976). The pace of life. *Nature, 259*, 557–559. https://doi.org/10.1038/259557a0.

Candia, C., Jara-Figueroa, C., Rodriguez-Sickert, C., Barabási, A.-L., & Hidalgo, C. A. (2018). The universal decay of collective memory and attention. *Nature Human Behaviour*(1).

Chadee, D., & Ditton, J. (2005). Fear of crime and the media: Assessing the lack of relationship. *Crime, Media, Culture, 1*(3), 322–332. https://doi.org/10.1177/1741659005057644.

Chermak, S. M., & Gruenewald, J. (2006). The media's coverage of domestic terrorism. *Justice Quarterly, 23*(4), 428–461. https://doi.org/10.1080/07418820600985305.

Coman, A. (2018). Predicting the decay of collective memory. *Nature Human Behaviour, 1*.

Curme, C., Zhuo, Y., Moat, H., & Preis, T. (2017). Quantifying the diversity of news around stock market moves. *Journal of Network Theory in Finance, 3*(1), https://doi.org/10.21314/JNTF.2017.027.

Cvetojevic, S., & Hochmair, H. H. (2018). Analyzing the spread of tweets in response to Paris attacks. *Computers, Environment and Urban Systems, 71*, 14–26. https://doi.org/10.1016/j.compenvurbsys.2018.03.010.

Dickinson, P. W. J. (1993). Fear of crime: Read all about it? the relationship between newspaper crime reporting and fear of crime. *British Journal of Criminology, 33*(1), 33–56. https://doi.org/10.1093/oxfordjournals.bjc.a048289.

Ditton, J., Chadee, D., Farrall, S., Gilchrist, E., & Bannister, J. (2004). From imitation to intimidation a note on the curious and changing relationship between the media, crime and fear of crime. *British Journal of Criminology, 44*(4), 595–610. https://doi.org/10.1093/bjc/azh028.

Ditton, J., & Duffy, J. (1983). Bias in the newspaper reporting of crime news. *British Journal of Criminology, 23*, 159. https://doi.org/10.1093/oxfordjournals.bjc.a047355.

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS ONE, 6*(12), 26752. https://doi.org/10.1371/journal.pone.0026752.

Downs, A. (1972). *Up and down with ecology: The issue-attention cycle.* The Public.

Fragkias, M., Lobo, J., Strumsky, D., & Seto, K. C. (2013). Does size matter? Scaling of CO2 emissions and US urban areas. *PLoS ONE, 8*(6), 64727.

Galtung, J., & Ruge, M. H. (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research, 2*(1), 64–90. https://doi.org/10.1177/002234336500200104.

Hall, C. M. (2002). Travel safety, terrorism and the media: The significance of the issue-attention cycle. *Current Issues in Tourism, 5*(5), 458–466. https://doi.org/10.1080/13683500208667935.

Harcup, T., & O'Neill, D. (2001). What is news? Galtung and Ruge revisited. *Journalism Studies, 2*(2), 261–280. https://doi.org/10.1080/14616700118449.

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communication, 18*(2), 40–60. https://doi.org/10.1111/jcc4.12001.

Joppke, C. (1991). Social movements during cycles of issue attention: The decline of the anti-nuclear energy movements in West Germany and the USA. *British Journal of Sociology,* 43–60. https://doi.org/10.2307/590834.

Kounadi, O., Lampoltshammer, T. J., Groff, E., Sitko, I., & Leitner, M. (2015). Exploring Twitter to analyze the public's reaction patterns to recently reported homicides in London. *PLoS ONE, 10*(3), 0121848. https://doi.org/10.1371/journal.pone.0121848.

Lane, J., & Meeker, J. W. (2003). Ethnicity, information sources, and fear of crime. *Deviant Behavior, 24*(1), 1–26. https://doi.org/10.1080/10639620390117165.

Leetaru, K., & Schrodt, P. A. (2013). Gdelt: Global data on events, location, and tone, 1979–2012. *ISA Annual Convention. Vol. 2. ISA Annual Convention* (pp. 1–49). Citeseer.

Liska, A. E., & Baccaglini, W. (1990). Feeling safe by comparison: Crime in the newspaper. *Social Problems, 37*, 360. https://doi.org/10.2307/800748.

Malleson, N., & Andresen, M. A. (2015). The impact of using social media data in crime rate calculations: Shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science, 42*(2), 112–121. https://doi.org/10.1080/15230406.2014.905756.

McDonald, S. (2009). Changing climate, changing minds: Applying the literature on media effects, public opinion, and the issue-attention cycle to increase public understanding of climate change. *International Journal of Sustainability Communication, 4*, 45–63.

Morley, D. (2003). *Television, audiences and cultural studies.* London, UK: Routledge.

Oliveira, M., Bastos-Filho, C., & Menezes, R. (2017). The scaling of crime concentration in cities. *PLoS ONE, 12*(8), 0183110.

Olteanu, A., Castillo, C., Diakopoulos, N., & Aberer, K. (2015). Comparing events coverage in online news and social media: The case of climate change. *ICWSM, 15*, 288–297.

Pak, A., & Paroubek, P. (2010). *Twitter as a corpus for sentiment analysis and opinion mining. LREC, Vol. 10*.

Peters, B. G., & Hogwood, B. W. (1985). In search of the issue-attention cycle. *The Journal of Politics, 47*(1), 238–253. https://doi.org/10.2307/2131074.

Petersen, K. K. (2009). Revisiting Downs' issue-attention cycle: International terrorism and US public opinion. *Journal of Strategic Security, 2*(4), 1. https://doi.org/10.5038/1944-0472.2.4.1.

PressGazette. NRS National press readership data: Telegraph overtakes guardian as most-read 'quality' title in print/online. (2018). https://www.pressgazette.co.uk/nrs-national-press-readership-data-telegraph-overtakes-guardian-as-most-read-quality-title-in-printonline/ Accessed 05/11/2018.

Prieto Curiel, R., Pappalardo, L., Gabrielli, L., & Bishop, S. R. (2018). Gravity and scaling laws of city to city migration. *PLoS ONE, 13*(7), 0199892.

Schläpfer, M., Bettencourt, L. M., Grauwin, S., Raschke, M., Claxton, R., Smoreda, Z., … Ratti, C. (2014). The scaling of human interactions with city size. *Journal of the Royal Society Interface, 11*(98), 20130789.

Schmidt, A., Ivanova, A., & Schäfer, M. S. (2013). Media attention for climate change around the world: A comparative analysis of newspaper coverage in 27 countries. *Global Environmental Change, 23*(5), 1233–1248. https://doi.org/10.1016/j.gloenvcha.2013.07.020.

Simon, H. A. (1971). *Designing organizations for an information-rich world*.

Sornette, D. (2009). Dragon-kings, black swans and the prediction of crises. *International Journal of Terraspace Science and Engineering. 2*(1), 1–18.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*(sup1), 234–240. https://doi.org/10.2307/143141.

Twitter. Twitter Developer Platform. (2018). https://developer.twitter.com Accessed 05/11/2018.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151.

Weng, L., Flammini, A., Vespignani, A., & Menczer, F. (2012). Competition among memes in a world with limited attention. *Scientific Reports, 2*(335).

Wu, F., & Huberman, B. A. (2007). Novelty and collective attention. *Proceedings of the National Academy of Sciences, 104*(45), 17599–17601.

Xu, W. W., Sang, Y., Blasiola, S., & Park, H. W. (2014). Predicting opinion leaders in Twitter activism networks: The case of the Wisconsin recall election. *American Behavioral Scientist, 58*(10), 1278–1293. https://doi.org/10.1177/0002764214527091.

Yang, J., & Leskovec, J. (2011). Patterns of temporal variation in online media. *Proceedings of the fourth ACM international conference on web search and data mining* (pp. 177–186). ACM.

Yuan, Y., Liu, Y., & Wei, G. (2017). Exploring inter-country connection in mass media: A case study of China. *Computers, Environment and Urban Systems, 62*, 86–96. https://doi.org/10.1016/j.compenvurbsys.2016.10.012.

Zelenkauskaite, A., & Balduccini, M. (2017). Information warfare and online news commenting: Analyzing forces of social influence through location-based commenting user typology. *Social Media + Society, 3*(3), 2056305117718468. https://doi.org/10.1177/2056305117718468.

Zhang, Z., Zhang, Y., Shen, D., & Zhang, W. (2018). The dynamic cross-correlations between mass media news, new media news, and stock returns. *Complexity*. https://doi.org/10.1155/2018/7619494.