Changing the Subject: Dynamic Discussion Monitoring in Twitter

Marçal Mora-Cantallops(✉) and Salvador Sánchez-Alonso

University of Alcalá, Alcalá de Henares, Madrid, Spain
{marcal.mora,salvador.sanchez}@uah.es

Abstract. In recent years, Twitter has become increasingly popular both as a social networking service (where users express their opinions) and as a tool for information retrieval. Many events that are commented, debated or argued online, however, are dynamic and unpredictable in nature, resulting in the need to derive the corresponding dynamic computational methodologies to track and to extract such changing topics, events and relevant content in a timely and unattended manner. In this paper, we propose a framework to accomplish two objectives: periodically obtaining the relevant topics of discussion among authorities and adapting the tracked keywords or hashtags accordingly to retrieve the most relevant possible information in each time window. The application of this framework to a case study reveals how our proposed approach is able to accurately and dynamically track the conversation around an event. Our method could therefore be applied to query Twitter in many domains, such as politics, sport events, marketing campaigns or media engagement, among others.

Keywords: Information retrieval · Twitter · Dynamic · Topic monitoring

1 Introduction

The adoption of Social Networking Services (SNS) (also referred as social media) has been growing steadily over the last years, with an estimated 3.5 billion active users and 45% penetration globally[1]. Many users have drifted away from mainstream media as its credibility scores have dropped across all venues and have embraced the new possibilities offered by the social media as a source of information, as news shared by contacts or friends seems to improve the level of trust; this effect is amplified when the contact sharing the story is perceived as an "opinion leader" [1].

Therefore, trends in social media have become highly relevant to understand public opinion and perception [2]. They can be useful to track public conversations and debates, to assess a product performance or to understand human

---

[1] https://wearesocial.com/global-digital-report-2019.

behavior, among many others. To be able to know where public attention is at a particular moment becomes a critical problem for researchers and practitioners alike as trending topics can both describe the opinion of a large community and provide the means to analyze it.

Not all SNS serve the same purpose, though. Twitter is a social networking service that has emerged as the reference in micro-blogging platforms [3]. It combines some unique features such as a very simple interface and a character limit in its postings, making it convenient for information retrieval and automatic processing purposes. Users post short messages (up to 280 characters currently) called tweets and are able to interact with the tweets posted by others (replying to them, quoting or retweeting them). Huge amounts of data are generated every day that can provide valuable information for many different domains, such as political communication, consumer behaviors, marketing campaigns and disaster management, among others. This data is freely and publicly available both through its site and its API (application programming interface), allowing for a near real-time monitoring of users' preferences, opinion and behavior. User messages often include hashtags as a way of explicitly marking the relevant topics, easing such monitoring and analysis.

User behavior and occurrences, however, are unpredictable and dynamic in nature. Not only many events can't be accurately predicted (e.g. natural disasters or accidents) but also it's even harder for researchers and analysts to anticipate the wording that will be used by users in their hashtags. It could be argued that a constant manual monitoring could mitigate such issues, adding any new relevant words to the tracking as soon as they are noticed. This approach, however, presents two main caveats:

– Late response, as there might be a (relevant) delay between the new topic emerging and the start of its tracking, losing what could potentially be relevant information or interactions.
– Human supervision, which might not be possible at all times (e.g. at night), resulting in either inefficient resource allocation or the inability to detect and track relevant changes at any time.

In this paper, we propose a dual framework to accomplish two objectives: periodically obtaining the relevant topics of discussion (in a predefined scope) while adapting the tracked hashtags accordingly in order to be able to retrieve the most relevant possible information in each time window in an unattended and timely manner, minimizing information loss and human intervention. Our framework uses two algorithms. A first algorithm is devoted to monitor a group of authorities; as a result, a list of hashtags under discussion is obtained. The second algorithm then uses the most relevant topics to reset the data stream filter, extracting and storing the related posts. This way, we maximize the probability of capturing and tracking the most relevant hashtags at a given point in time.

The remainder of this paper is organized as follows: related works are outlined in the following section. The proposed approach is described in detail in the methodology section, while a test case study is presented in Sect. 4. Finally, conclusions are discussed in Sect. 5.

2   Background

2.1   Social Media

Social media (or SNS) are web-based services that allow individuals to create a profile, articulate a list of connections to other users, view and traverse such connections between users, and share content [4]. SNS users tend to interact with other users to whom they already have some kind of social tie; therefore, online conversation through these platforms may more closely resemble opportunities for everyday conversation [5] about any topic than a more structured online forum where most users are strangers to each other. According to previous research, young people already show a clear preference for online engagement and organization; politically engaged young people integrate social media use into their existing organizations and political communications [6]. Many works have studied the effect of social media, specially Twitter, as a facilitator in political campaigns [7] and protests [8] worldwide.

One important aspect to consider is how information spreads in online social networks. Mønsted et al. [9] found that the best explanation was a complex contagion model, implying that information diffusion is affected not only by the number of exposures to a piece of information, but also by the exposure to multiple sources and their social influence. Users tend to follow other users that become their sources in this media environment; those perceived as leaders of opinion become more influential and the information diffused by them, more "viral" [1]. Thus, the dominant mode of information acquisition is through "incidental news", "news content encountered on mobile devices while visiting social media sites, in a process that is derivative of social media interactions rather than deliberately sought for" [10].

Therefore, while social media is rapidly becoming the main source of information for many citizens, the spread of information is irregular, difficult to predict in nature and incidental. Methods that aim to track and retrieve such conversations are required to be flexible and dynamic enough to be able to follow them and minimize information loss.

2.2   Topic Tracking

To be able to properly track opinion over time has been one of the main concerns by the public opinion analysts for a long time [11]. With the advent of Twitter, public opinion can be tracked continuously and in real time. The high volume of information and the myriad different topics that are being talked in any particular time, however, stand in the way of direct tracking. One way to overcome such limitation is to look at the Twitter hashtags (keywords or terms starting with "#"), which are the most common feature for users to connect and relate to within a larger networked discourse [12]. Posts that contain hashtags are prone to contain more informational value than non-hashtagged tweets, as they are related with longer messages, additional hashtags and hyperlinks [13]. Moreover, hashtag use is not limited to one's own network, as Enli and Simonsen

found; journalists and politicians, for example, use them to reach outside their personal networks [14], demonstrating that their use of social media is closely connected to their professional practice.

Accurate and appropriate hashtag tracking is, therefore, crucial for discourse and conversation analysis in Twitter. But, in terms of information life cycle, it is difficult to predict what hashtag is going to be adopted by the community in a given discussion and whether it will alternate with custom or modified tags, so dynamic tracking is also key. In spite of this, most studies opt for a static approach. To cite a few examples, in [15], the authors used Twitter data to monitor a constitutional referendum in Italy. Even though they recorded tweets during five weeks, the extracted keywords were static: five hashtags that were manually selected by the authors. Another similar case is found in [16], where #WorldEnvironmentDay is tracked to understand public opinion about the subject, but user variants that might be widely used or hashtags in other languages were discarded. When Takahashi et al. [17] analyzed communication on Twitter during a natural disaster (a typhoon), they used four fixed hashtags, even though the typhoon lasted for five days. Missing data and losing track of the conversation when hashtags remain static is inevitable, as Tsakalidis et al. [18] acknowledge in their work about the EU 2014 election trends, even when they "aggregated tweets written in the respective language that contained a party's name, its abbreviation, its Twitter account name and some possible misspells" and "excluded several ambiguous keywords in an attempt to reduce the noise". In all cases, major candidate or party names along with generic terms were used, leaving little space to unusual terms, unexpected events or minor candidates, losing potentially relevant information. It is also worth noting how different languages, abbreviations, misspells or ambiguous keywords are problematic.

A few works have addressed real-time Twitter analysis with topic detection. Choi and Park [19] proposed a method to detect emerging topics on Twitter using High Utility Pattern Mining (HUPM), which takes frequency of appearance and utility of words into account. Although their method works well to detect topics in known datasets, it is not designed to dynamically use the resulting topics for extraction. A similar issue appears in the work of Adedoyin-Olowe et al. [20]; their work aims to detect relevant events from a collection of Tweets but, again, this is done in post-processing. There is no live adaptation where, for example, an event is detected and immediately followed by the Tweet extractor. This dynamic adaptation is found in Gaglio et al. [21], who proposed a system able to progressively refine its query to include new relevant terms, reflecting the emergence of new topics of trends. In their conclusions, they also noted how "other systems were unable to capture the social aspects of the observed events [...] every time the users left the main topic and started to talk about unexpected events". Their work, however, presents another limitation that might be relevant in some contexts: it requires an initial set of keywords to track and term selection depends on the collected tweets, not taking into account the authority of the users tweeting them. Our proposed solution will aim to overcome this situation.
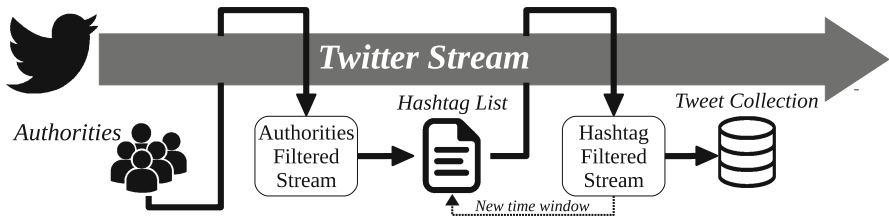
Fig. 1. Twitter is monitored using a initial set of accounts that are considered "authorities" in the topic of discussion or analysis. The acquired tweets are analyzed to extract the relevant terms. Terms are weighted, so for every time period an ordered list of topics of discussion is obtained, sorted by relevance. At the start of each time window, the N top terms are passed into the extraction script, updating the applied filter into the second stream. The collected tweets are stored for later analysis.

3   Methodology

The proposed solution is intended to be used as an automatic and unattended tool for tracking the conversation that is generated around a particular area or topic, although manual tuning is still possible at any moment in case running changes are needed. A general overview of the process is represented in Fig. 1, following four main steps, namely: (1) Selection of authorities, (2) Monitoring authorities, (3) Hashtag selection and (4) Conversation tracking.

3.1   Selection of Authorities

Users that, according to a certain criteria defined by the researcher, lead the conversation or are relevant in the topic of analysis are considered "authorities". To study the debate of news stories, for example, one would follow the media, the experts and the influential users in that matter. In a sport event, the list of authorities could be composed by not only by the team's accounts but also by relevant journalists and analysts.

Having a good list of authorities is a prerequisite for the proposed approach. This list will be followed using the Twitter API and will prevent the topic tracking from drifting to potentially undesired scopes.

3.2   Monitoring Authorities

Once a list of authorities is defined, the Twitter API stream can be filtered by following these users. In particular, the stream will contain and deliver posts that are either created or retweeted by any of the authorities, are replies to any Tweet created by any of the authorities or are retweets of any of their Tweets.

Every time a Tweet matching the filter is received, the script will proceed to process it, as described in Algorithm 1. This script has two main purposes: to generate the weighted collection of hashtags and to control whether a new time window has started.

Algorithm 1. Tweet Monitor

Input: A Tweet delivered from the authorities stream (matching the followed
authorities) **and** $C_t \longleftarrow$ Hashtag collection for current time period t
Output: $C_t' \longleftarrow$ Updated collection of hashtags for time window t (if still in
the same time window) **or** a new collection of hashtags for time
window $t+1$ (C_{t+1}) (in case it is a new time window)
begin

```
    /* Check whether it is a new time window                    */
```
    if *new time window* **then**
```
        Save C_t to file              /* Store C_t for later processing */
        Clear C_t                              /* C_t is now C_{t+1} */
```
```
    /* Process the received Tweet from the stream               */
```
    for *each Hashtag* **in** *Tweet* **do**
        if *Hashtag* **not in** C_t **then**
            Add Hashtag to C_t with weight w
        else
            Increase $C_t[Hashtag]$ w units

The collection of hashtags C_t is, in essence, an ordered list of hashtags found on all the posts recovered during a single time window. For each Tweet received, the script extracts all its hashtags (identified by their # symbol). Each hashtag can be assigned a weight w; the weight function may be, again, defined by the user. Depending on the intent, the weight function could take many forms. For instance, one might decide to simply count the number of times a hashtag appears, but it is also possible to weight a hashtag according to the posting user number of followers (assuming that it might have more impact or relevancy), the logarithm of that number (to avoid being too influenced by dominant users) or more complex functions.

The time window t must be defined by the user and it allows to adapt the monitoring to different paces. A long range tracking (a marketing campaign, for instance) might require a hourly or daily tracking, while shorter events (e.g. an sports event or a natural disaster) could need a window of only a few minutes. Regardless of the time window used, the Tweet Monitor script saves the collection of hashtags for each time period.

3.3   Hashtag Selection

Before proceeding to monitor the hashtags used by the authorities, a selection step is required. Here, an approach mixing "hot" and "emerging" topics is suggested.

To obtain the list of hashtags to be tracked to monitor the derived conversation in Twitter, the following steps are completed:

1. If desired, a list of stop words can be added to remove common everyday spurious hashtags (such as #goodmorning or #happymonday).

2. Every time period, the script looks at the hashtag collection for that window and extracts the top M hashtags in absolute terms by their weight. These are supposed to be the topics that are being strongly discussed and posted by the authorities.
3. Every time period, the hashtags for the corresponding window (C_t) are compared to the previous window C_{t-1} in order to obtain the E hashtags with the biggest increase (note that this might be extended up to the n previous periods). These are potential emerging topics, that might not yet be on the peak of the discussion but are rapidly growing.
4. The top M and E hashtags are combined into the final list, consisting of N (at maximum $M + E$) hashtags, as represented in Fig. 2.
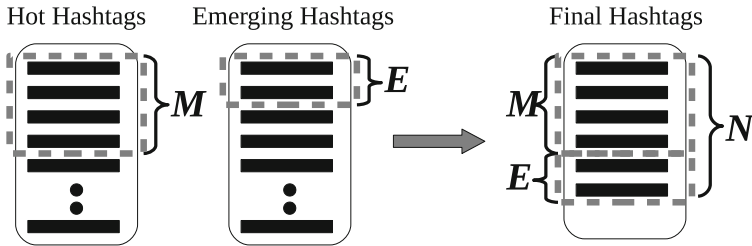

Fig. 2. The list with the top N hashtags combines the biggest terms of discussion (M) with the fastest growing ones (E). For each case, M and E can be adapted to the researcher needs.

3.4   Conversation Tracking

Once the final list of hashtags is obtained, a separate stream can be set to track such hashtags. The Twitter API stream, in this case, delivers all Tweets (from the provided stream, which is in turn a sample of all public Tweets) that match any of that hashtags in their texts (and ignoring case). These Tweets can be processed or stored for a later analysis. Tracked hashtags are dynamically changed in each time period; they are renewed with a new list coming from the previous step, allowing for the unattended and dynamic tracking of the topics present in the authorities discussion.

4   Experimental Results

4.1   Case Study

In order to analyze and provide a few insights on the proposed methodology, a case study has been used. In particular, the conversation around the 2019 Eurovision Song Contest was tracked during the event, between 20:00 CEST on the 18th of May and 02:00 on the following day. As known, Eurovision is an event that is broadcast internationally and attracts a large number of interactions in Twitter, as "users offer their own running commentary on the universally shared media text of the event as it unfolds live" [22]. For this case, the weight function

w was defined as the number of followers that the tweeting user had (therefore, the more followers, the more relevance in the discussion) and the time window t was set at two minutes (given that each performance lasts for 3–4 min).

The following subsections will focus on the hashtag collections generated by the Tweet Monitor; the selection of the top N hashtags and their extraction is considered straightforward after that step. A N of 25, with 20 hot (M) and 5 emerging (E) hashtags will be assumed for this case.

4.2   Authorities

For this case study, the official Eurovision Twitter accounts of every participating country's public broadcaster were used as authorities. Although there were up to 41 participant countries (26 of which performed at the final event), only active and verified accounts specific to Eurovision were included. The selection included a total of 20 accounts, composed by the official Eurovision account (@eurovision) and 19 official accounts managed by each country's broadcasting company, where available. These include, among others, @kaneurovision from Israel, @bbceurovision from the UK, @sbseurovision for Australia, @yleeurovision for Finland and @eurovision_tve for Spain.
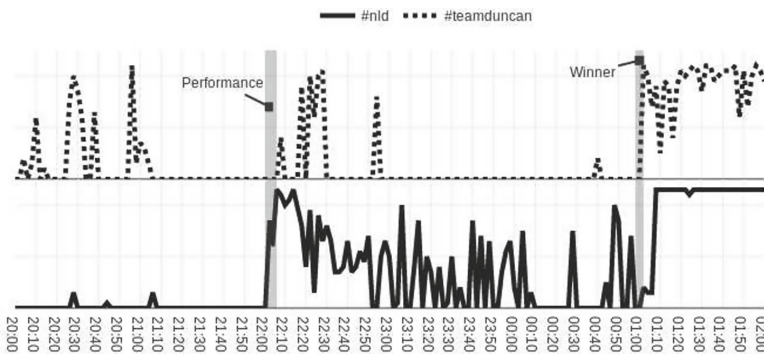

Fig. 3. Ranked position of each two hashtags supporting the same performance (#nld and #teamduncan) in each time window

4.3   Results

A total of 1814 different hashtags were obtained during the six hour extraction. After extracting the top 25 hashtags in each time period, 429 unique hashtags remain. Each hashtag is then ranked according to its weight during the time window. The first visible result is that, although the official hashtags (#eurovision, #daretodream) are consistently in the top 3 (and, therefore, constantly tracked) there are other variations that are also often present in the conversation, such as #esc2019 or #eurovision2019.

Unpredictable Topics. As previously noted, most studies start by monitoring a fixed list of hashtags. That list may include all expected words (in this case, the official hashtags, both for the festival and for each country). Therefore, for a country like The Netherlands, which was the contest winner, one could have expected to track #nld and, maybe, the name of their song (#arcade). But the fan community adopted a different hashtag, #teamduncan, composed with the singer's (Duncan Laurence) first name. As can be seen in Fig. 3, this hashtag is highly relevant not only right after the performance, but also before the contest started (when the official hashtag was not online yet) and after the winner was announced, responding more quickly than the conventional tag.
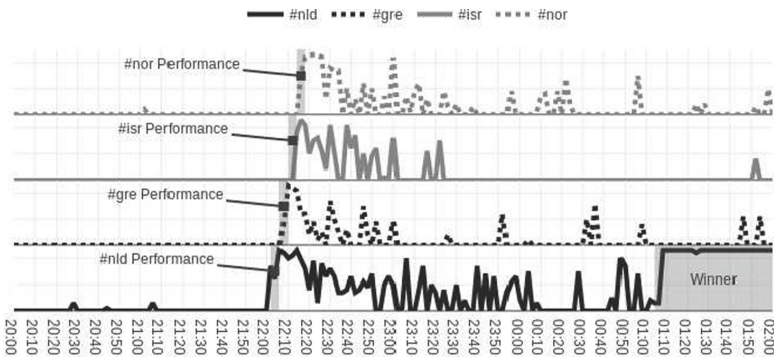

Fig. 4. Ranked position (in the top 25) of each hashtag related to four consecutive performances (#nld, #gre, #isr and #nor) in each time window

Events. On one hand, to obtain a more efficient extraction, it is interesting to track each hashtag only when it is relevant enough. On the other, some events might happen spontaneously or are hard to predict. Figure 4 is a good example of this. First, it must be noted that these four consecutive performances by as many countries get into the top hashtags right in the moment (and not before) when the authorities start talking about them during their respective performances. Second, some events (such as the performance by #isr) reach a peak when they happen, are on the conversation for a while, decay and disappear from the list (note that Israel was not a favorite in this edition). This also shows how irrelevant topics are "forgotten" promptly. Third, the case of the #nld, one of the favorites and final winner, where discussion lasts for longer, then goes down during the voting and goes back to the top when the winner is announced.

Languages. This method is also able to track variations of the same topic in different languages (in case they are relevant enough). In the case of Eurovision, the Polish word #eurowizja was also among the top 25 hashtags in 43 out of 181 time windows. Other example can be found in the calls for boycott (the festival took place in Israel, raising controversy due to the situation with Palestine) with #boicoteurovision2019 and #boycotteurovision2019 showing similar levels of usage.
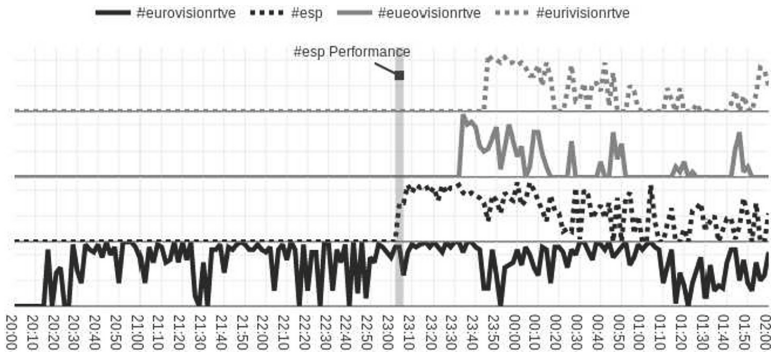
Fig. 5. Ranked position of each hashtag related to #eurovisionrtve among the top 25 in each time window

Misspellings or Alternatives. Alternative words for the same concepts are also found in a good amount. Spanish fans, for example, used #esp, #spain or even #lavenda (name of the song) to talk about their song. In a similar way, Swedish fans used #swe or #sweden interchangeably. Not only some hashtags are unpredictable, but their alternatives or misspellings created by users might also be problematic. A paradigmatic case can be seen in Fig. 5, where the hashtags related to the Spanish Public Television (RTVE). The official hashtag, #eurovisionrtve, is prominent over the whole festival, often among the top three. However, after the Spanish performance, two misspelled related hashtags appear in the conversation: #eueovisionrtve and #eurivisionrtve (note how the "e" is next to the "r" and the "i" is next to the "o" in the Spanish keyboard). Although some misspellings might be anecdotal, both these hashtags were relevant (among the top 25) in many time windows until the end of the festival; not tracking them would derive in a definite loss of information.

5   Conclusions

Even relatively simple events such as the Eurovision Song Contest derive in complex conversations. In our study, more than 1800 different hashtags were generated by the authorities (official and verified accounts) in the subject, with 429 unique different hashtags among the top 25 topics in at least one time window. Previous studies analyzed Eurovision using only a few static hashtags; [22], for instance, tracked only three keywords (#eurovision, #esc and #sbseurovision), missing interesting information only traceable through misspelled hashtags or user generated tags that must be detected in real time. We have shown how our unattended approach could enrich this kind of extractions, adding a dynamic tracking of the conversation among the relevant accounts to capture the resulting public discussion. In particular, it should be noticed how #esc was not even relevant as a hashtag for the 2019 edition, as users ended using #esc2019 instead.

Thus, the proposed methodology is able to track hashtags in an unattended and dynamic manner, capturing the hashtags used by the authorities of the

community of interest and adapting them as the conversation changes or moves, even if dramatically, while being robust against disruptions and avoiding the requirement of an initial set of words; term selection depends exclusively on the terms used by the authorities and is also able to take their relevance into account.

This method is limited by a few restrictions, however. First, it requires a fairly accurate identification of the authorities, users that are relevant in the discussion. In political contexts or in many events, they should be arguably straightforward to identify (political leaders, influencers, etc.) but there might be some scopes where this leadership is less clear. Second, the determination of the time window and the weight function is also important for its behavior; they should be determined for each case using proper testing, which is not always possible as the event might be unique. In any case, similar events might be useful to tune the algorithm.

Future work will look at how to dynamically adapt the list of authorities, as not all the users driving the conversation might be included in the initial list; in this way, it would be possible to add new members or replace the less active ones after identifying the most influential users in a particular time window [23]. Another venue for further research is appropriate parameter setting in each situation, as to be able to correctly set the time window, the weight function and the number of hashtags to be followed is also critical for the method's generalization.

References

1. Turcotte, J., York, C., Irving, J., Scholl, R.M., Pingree, R.J.: News recommendations from social media opinion leaders: effects on media trust and information seeking. J. Comput. Commun. **20**, 520–535 (2015). https://doi.org/10.1111/jcc4.12127
2. Fang, Y., Chen, X., Song, Z., Wang, T., Cao, Y.: Modelling propagation of public opinions on microblogging big data using sentiment analysis and compartmental models. Int. J. Semant. Web Inf. Syst. **13**, 11–27 (2017). https://doi.org/10.4018/ijswis.2017010102
3. Gadek, G., et al.: Topological and topical characterisation of Twitter user communities. Data Technol. Appl. **52**, 482–501 (2018). https://doi.org/10.1108/DTA-01-2018-0006
4. Boyd, D.M., Ellison, N.B.: Social network sites: definition, history, and scholarship. J. Comput. Commun. **13**, 210–230 (2007). https://doi.org/10.1111/j.1083-6101.2007.00393.x
5. Hampton, K.N., Shin, I., Lu, W.: Social media and political discussion: when online presence silences offline conversation. Inf. Commun. Soc. **20**, 1090–1107 (2017). https://doi.org/10.1080/1369118X.2016.1218526
6. Vromen, A., Xenos, M.A., Loader, B.: Young people, social media and connective action: from organisational maintenance to everyday political talk. J. Youth Stud. **18**, 80–100 (2015). https://doi.org/10.1080/13676261.2014.933198
7. McGregor, S.C., Mourão, R.R., Molyneux, L.: Twitter as a tool for and object of political and electoral activity: Considering electoral context and variance among actors. J. Inf. Technol. Polit. **14**, 154–167 (2017). https://doi.org/10.1080/19331681.2017.1308289

8. Jost, J.T., et al.: How social media facilitates political protest: information, motivation, and social networks. Polit. Psychol. **39**, 85–118 (2018). https://doi.org/10.1111/pops.12478

9. Mønsted, B., Sapieżyński, P., Ferrara, E., Lehmann, S.: Evidence of complex contagion of information in social media: an experiment using Twitter bots. PLoS ONE **12** (2017). https://doi.org/10.1371/journal.pone.0184148

10. Boczkowski, P., Mitchelstein, E., Matassi, M.: Incidental news: how young people consume news on social media. In: Proceedings of the Hawaii International Conference on System Sciences, pp. 1785–1792 (2017). https://doi.org/10.24251/hicss.2017.217

11. Green, D.P., Gerber, A.S., Boef, S.L. De.: Tracking opinion over time: a method for reducing sampling error. Public Opin. Q. **63**, 178–192 (2002). https://doi.org/10.1086/297710

12. Bruns, A. and Burgess, J.: Twitter hashtags from ad hoc to calculated publics. Hashtag publics: the power and politics of discursive networks, pp. 13–28 (2015)

13. D'heer, E., Verdergem, P., De Grove, F.: #MissingData: a methodological inquiry of the hashtag to collect data from Twitter. Selected Papers AoIR 2016 17th Annual Conference of the Association of Internet Research (2016)

14. Enli, G., Simonsen, C.A.: 'Social media logic' meets professional norms: Twitter hashtags usage by journalists and politicians. Inf. Commun. Soc. **21**, 1081–1096 (2018). https://doi.org/10.1080/1369118X.2017.1301515

15. Fano, S., Slanzi, D.: Using Twitter data to monitor political campaigns and predict election results. In: De la Prieta, F., et al. (eds.) PAAMS 2017. AISC, vol. 619, pp. 191–197. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-61578-3_19

16. Reyes-Menendez, A., Saura, J.R., Alvarez-Alonso, C.: Understanding #worldenvironmentday user opinions in Twitter: a topic-based sentiment analysis approach. Int. J. Environ. Res. Public Health. **15**, 2537 (2018). https://doi.org/10.3390/ijerph15112537

17. Takahashi, B., Tandoc, E.C., Carmichael, C.: Communicating on Twitter during a disaster: an analysis of tweets during Typhoon Haiyan in the philippines. Comput. Human Behav. **50**, 392–398 (2015). https://doi.org/10.1016/J.CHB.2015.04.020

18. Tsakalidis, A., Papadopoulos, S., Cristea, A.I., Kompatsiaris, Y.: Predicting elections for multiple countries using Twitter and polls. IEEE Intell. Syst. **30**, 10–17 (2015). https://doi.org/10.1109/MIS.2015.17

19. Choi, H.J., Park, C.H.: Emerging topic detection in Twitter stream based on high utility pattern mining. Expert Syst. Appl. **115**, 27–36 (2019). https://doi.org/10.1016/j.eswa.2018.07.051

20. Adedoyin-Olowe, M., Gaber, M.M., Dancausa, C.M., Stahl, F., Gomes, J.B.: A rule dynamics approach to event detection in Twitter with its application to sports and politics. Expert Syst. Appl. **55**, 351–360 (2016). https://doi.org/10.1016/j.eswa.2016.02.028

21. Gaglio, S., Lo Re, G., Morana, M.: A framework for real-time Twitter data analysis. Comput. Commun. **73**, 236–242 (2016). https://doi.org/10.1016/j.comcom.2015.09.021

22. Highfield, T., Harrington, S., Bruns, A.: Twitter as a technology for audiencing and fandom. Inf. Commun. Soc. **16**, 315–339 (2013). https://doi.org/10.1080/1369118x.2012.756053

23. Mahmoudi, A., Yaakub, M., Abu Bakar, A.: New time-based model to identify the influential users in online social networks. Data Technol. Appl. **52**, 278–290 (2018). https://doi.org/10.1108/DTA-08-2017-0056