Any Publicity is Good Publicity: Positive, Negative and Neutral Tweets Can All Become Trends

Marco A. Palomino
School of Engineering Computing and Mathematics
University of Plymouth
Plymouth, UK
marco.palomino@plymouth.ac.uk

Aditya Padmanabhan Varma*
School of Electronics Engineering
Vellore Institute of Technology
Chennai, India
adityapvarma@gmail.com

***Abstract*—Anecdotal evidence suggests that Twitter trends are characterised by highly polarised tweets. However, specific experiments intended to measure the correlation between the emergence of a trend and the overall sentiment expressed on it have been few and limited. Thus, we have launched an investigation to ascertain the nature of the relationship between trends and strength of sentiment. As a testbed for our experimentation, we have retrieved a large collection of tweets related to the COVID-19 pandemic, in the particular context of the UK Government briefings broadcasted in the media. Our results highlight the presence of a significant percentage of trends with a nearly neutral sentiment. Indeed, there does not seem to be an apparent correlation between trends and polarity.**

***Index Terms*—Social media, Twitter, sentiment analysis, trending topics, SentiStrength, TextBlob**

I. Introduction

Although popular wisdom suggests that "any publicity is good publicity", evidence from previous research shows varying results. While some studies maintain that positive publicity help retail marketers to ensure consumers remain loyal towards certain brands [1], others argue that negative publicity can increase sales, especially when a product is relatively unknown [2]. For example, looking at 240 fiction book titles reviewed by *The New York Times*, Berger *et al.* found that negative reviews for books by established authors led to a 15% decrease in sales; yet, for books by relatively unknown authors, negative publicity increased sales by a significant 45%. Follow-up studies have confirmed that bad reviews draw attention to works that otherwise would have been unnoticed [2].

We are particularly interested in online opinions and the case for *sentiment analysis*, which has emerged as an alternative to categorise opinions automatically [3], [4]. Social networking platforms, such as *Twitter*, are now full of reviews, praise and criticism [5]. Businesses stay in touch with their customers through Twitter and give them information about deals and brands [6]. Governments make use of Twitter to disseminate campaigns—for example, health campaigns [7]—and individuals have employed the platform to promote their careers, services and partnerships [8]. Undoubtedly, Twitter is a very appealing source of data to experiment with opinion mining and sentiment analysis [9].

*This work was carried out while Aditya Padmanabhan Varma was enrolled on the University of Plymouth's SECaM Semester Abroad Scheme.

In order to index the vast amount of data available on Twitter—500 million tweets are published daily [10]—computer scientists have developed algorithms to capture *trending topics*, or simply *trends*—i.e., themes that experience a surge in popularity [11]. Such trends typically refer to terms and phrases that reflect current events—for instance, `#quarantine`, which refers to one of the health measures set in response to the 2019 coronavirus outbreak—and often include keywords and phrases extracted from popular conversations—for instance, `#StayHomeSaveLives`, which refers to a recommendation included in the UK Government's information campaign introduced in March 2020.

Detecting trends in Twitter is challenging, because it involves a huge amount of data. Several natural language processing (NLP) methods have been developed for this purpose—our previous research has approached some of them, see Palomino and Murali 2019 [12]. However, our goal is to investigate how sentiment analysis can aid in the discovery of trends: we want to ascertain if trends are characterised by a strong sentiment. To the best of our knowledge, no study has yet addressed this issue in detail. Casual observations may support different arguments to explain the correlation between trends and sentiment. However, we want to prove if it is correct to assume that trends are characterised by a highly polarised sentiment. There is a need for a study that looks into this subject in depth, and this is the motivation behind this paper.

The remainder of this paper is organised as follows: We begin by reviewing the related work, and then we describe the corpus we used for our experiments. Afterwards, we explain how we identify trends and we discuss the results of evaluating the sentiment of the trends we found in our experimental corpus. Finally, we offer our conclusions.

II. Related Work

Social media has influenced the lives of most adults in the UK since its inception. The UK Office for National Statistics found that 66% of the UK population regularly used social media in 2017, rising from 45% in 2011 [13]. While social media users are not consolidated on a single platform, Twitter has some distinct advantages when compared with others. For instance, the amount of publicly available data on Facebook is small, whereas nearly all the data on Twitter is public.

Twitter is a microblogging platform that enables people to post short messages—*tweets*—expressing what they are willing to share [14]. Commonly, Twitter users employ *hashtags*—words or phrases preceded by a hash sign '#'—to categorise tweets topically, so that people can follow conversations centring on a specific topic. A detailed description of Twitter has been published by Murthy [8].

As the research literature involving Twitter keeps growing, it has become clear that tweets contain valuable information for public health research [15]. However, the mining of tweets related to public health has concentrated on *syndromic surveillance* [16], which tracks trends in medical conditions over time. Indeed, Lampos and Cristianini [17] and Culotta [18] correlated tweets mentioning the flu and related symptoms with historical data; De Quincey and Kostkova collected tweets to monitor the H1N1 pandemic [19]; and Scanfeld *et al.* evaluated the public understanding of antibiotics by manually reviewing tweets about incorrect antibiotic use [20].

Public health research would benefit greatly from the research on Twitter trends. The discovery of a new trend can be an (early) "symptom" of changes in population health [21]. If we had robust metrics which reflected the emergence of trends and their links to available information in real-time, all kinds of useful public health applications could be developed.

Twitter has already got an algorithm to identify trends [22]; yet, the algorithm is tailored for individual users [23]. Thus, the trends determined are based on the location of each specific user and the accounts she follows, and do not contemplate the wider context, which is critical in public health.

Former research on trends started as an attempt to index Twitter content. For example, *TweetMotif* [24], a service that summarised tweets and provided an overview of what people were discussing, was used to discover trends. Although TweetMotif is no longer operational, detecting trends has remained an active subject of investigation [25]–[29]. Benhardus and Kalita [23], for instance, used *TF-IDF* [30], combined with a number of heuristics, to assign weights to the different terms comprised in a tweet—the terms with the greatest weights were used to identify trends. TF-IDF, which is short for *term frequency – inverse document frequency*, is a numerical statistic that is intended to reflect how important a word is to a document in a corpus [31]—in our case, a document is a tweet. We have also incorporated TF-IDF into our work, as we will discuss later in Section IV. However, we do not experiment with different heuristics, because measuring the strength of the sentiment expressed in a trend is more important to our work than the performance of a particular algorithm.

Shamma *et al.* [32] looked into ongoing temporal conversations to find *peaky* topics—topics that show highly localised, temporary interest. Although Shamma *et al.* mined the text across tweets, they did not consider the correlation between strength of sentiment and trends, which is central to our work.

Our interest in sentiment derives from the recent developments in the computational study of opinions and subjectivity in text [33]. Plenty of applications require the analysis of emotions expressed in opinions as part of their operation.

Sentiment analysis aims to systematically identify, extract, quantify, and study opinions about specific topics, and attitudes towards particular entities [34], [35]. Sentiment analysis has a great potential as a technology to enhance the capabilities of customer relationship management and recommendation systems—for example, showing which features customers are particularly happy about, or excluding from recommendations items that have received negative feedback [36].

The basic tasks of sentiment analysis are *emotion recognition* [37] and *polarity detection* [38]. While the first task focuses on identifying a variety of emotional states, such as "anger", "sadness" and "happiness", the second one is either a binary classification task—whose outputs are 'positive' versus 'negative', 'thumbs up' versus 'thumbs down', or 'like' versus 'dislike'—or a ternary classification task—whose outputs are 'positive', 'neutral' or 'negative'. Several sentiment analysis tools have been developed lately—both Feldman [34] and Ribeiro *et al.* [39] claim that 7,000 articles on sentiment analysis had been written up by 2016, while dozens of start-ups are developing sentiment analysis solutions.

For the purpose of our work, we have selected two specific tools: *SentiStrength* [40] and *TextBlob* [41]. The reason why we have chosen two separate tools is the lack of consensus among them [42]. Hence, rather than relying on a single tool, we prefer to consider a couple of them and compare and contrast their differences.

SentiStrength was specifically implemented to determine the strength of sentiment in informal English text, using methods to exploit the de-facto grammars and spelling styles of the informal communication that regularly takes place in social media, blogs and discussion forums [43].

TextBlob is a Python library which offers an API to perform NLP tasks, such as noun phrase extraction, language translation and spelling correction [41]. With respect to sentiment analysis, TextBlob provides two options for polarity detection: `PatternAnalyzer`, which is based on the data mining *Pattern* library developed by the Centre for Computational Linguistics and Psycholinguistics (CLiPS) [44], and a `NaiveBayesAnalyzer` classifier, which is a *Natural Language ToolKit* (NLTK) classifier trained on movie reviews [45], [46]. The default option for sentiment analysis is `PatternAnalyzer`, and that is precisely the option we favoured, because we are not working with movie reviews.

We will now explain how the corpus we used for our experiments was retrieved, and we will describe its features.

III. Experimental Corpus

As a testbed for our experiments, we gathered a corpus of 409,761 tweets about COVID-19 on 22 April 2020. We chose this particular date, because it was when the UK Foreign Secretary, Dominic Raab, delivered a press briefing to address the UK Government's response to the situation. Daily press briefings were held by the UK Government between 16 March 2020 and 23 June 2020. All the slides, datasets and transcripts employed in the press briefings are available for free on the UK public sector information website: `GOV.UK` [47].

The British press started to cover the news about a COVID-19 vaccine at the start of April 2020, when the first human trials began in Europe [48]. A significant investment was made on these trials, and we assumed the Foreign Secretary's briefing on 22 April 2020 would address this topic and spark off the discussion on Twitter. Thus, we thought this would be an ideal opportunity to capture tweets with a strong sentiment attached to them, either in the form of Government's criticism or concern for the prevailing situation.

We expected the briefing on 22 April 2020 to begin at around 16:30, and we decided to start the retrieval of tweets a couple of hours prior to the beginning of the briefing, and kept the retrieval process going for a couple of hours after the end of the briefing. To be precise, we captured our first tweet at 14:24:39, and the last one at 18:56:27, spanning a total duration of 4 hours, 31 minutes and 48 seconds.

We retrieved our tweets using *Tweepy* [49], an open-source, Python library for retrieving tweets in real time. Tweepy makes it easier for us to interact with all the Twitter RESTful API methods—including the *Twitter Streaming API*—by handling authentication and connection [50], [51]. The program in charge of gathering tweets was executed in Plymouth (UK).

To ensure we were actually capturing information about COVID-19, we looked specifically for tweets comprising the hashtags listed in Table I. Note that Table I also displays the number of tweets retrieved for each hashtag. The figures reported on Table I do not sum to give the total number of tweets available in the corpus: 409,761. This is because there are many tweets which include two or more of the hashtags listed in Table I. Also, the text of some of the tweets in the corpus may not include explicit occurrences of the hashtags listed in Table I; yet, the Streaming API would provide us with such tweets if the hashtags appeared as part of URLs or metadata associated with those tweets [51].

TABLE I
HASHTAGS AND NUMBER OF TWEETS PER HASHTAG.

Hashtag	Number of tweets
#covid19	238,432
#coronavirus	116,557
#stayhome	31,820
#covid_19	11,068
#socialdistancing	6,510
#covid-19	4,636
#covid2019	2,341
#flattenthecurve	2,124
#coronavirusoutbreak	2,058
#sarscov2	1,861
#virus	1,211

Figure 1 shows the number of tweets we retrieved every 30 minutes. On average, we retrieved 81,952 tweets per hour between 14:24 and 19:24; yet, we retrieved more than 90,000 tweets per hour for the first three hours. During the 10-minute period between 16:20 and 16:30, which was just before the beginning of the briefing, we retrieved 16,043 tweets.

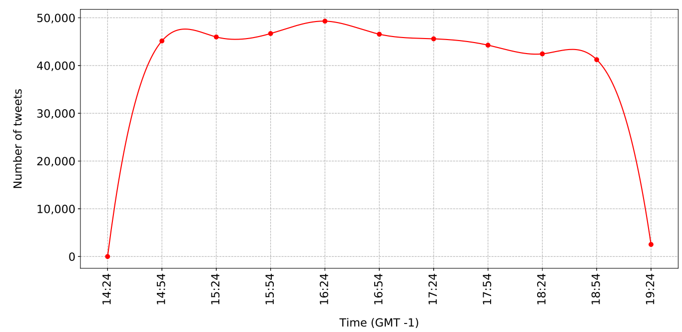

Fig. 1. Distribution of tweets over the retrieval.

Each tweet was retrieved as a `status object`—in the context of Twitter development tools, tweets are known as *status updates* [52]. The Streaming API provided the tweets and their corresponding metadata in Java Script Object Notation (JSON) format, and we produced a Python parser to extract the text of the tweets and other relevant information, such as the time when the tweets were published and the identifiers of the users who published those tweets. To archive the tweets we collected, we uploaded them into a MySQL database, which we later used to study the corpus.

Although we retrieved a corpus about COVID-19, we do not intend to monitor the spread of the disease or advise public health interventions. We chose COVID-19, because it is a key topic, which is likely to encourage the support or opposition of a large number of people. Thus, our corpus will allow us to find trends and assess the sentiment expressed on them.

IV. TREND DISCOVERY

Our definition of trend is a modified version of the one proposed by Benhardus and Kalita [23]: *A trend is a word, or combination of words, that experiences an increase in usage, both in relation to its long-term usage and in relation to the usage of other words*. Our system to identify trends is written in Python and runs under *Google Colaboratory*, or *Colab*, Google's environment for interactive development [53].

Ideally, we would like to process tweets in real-time, and cluster them as they are being retrieved. However, for experimentation purposes, we opted for retrieving the tweets first and processing them later, which gave us the additional advantage of testing different processing approaches.

Pre-processing, the practice of "cleaning" and preparing text for its analysis, is critical to trend discovery. Lack of pre-processing often results in lengthy computation and inaccurate results. Therefore, for each tweet in the corpus, we pre-process its text as follows: we change it entirely to lowercase—for ease of comparison—and remove unnecessary whitespaces, carriage return and newline characters, emoticons and URLs. Next, we tokenise the remaining "clean" text and remove stop words. Figure 2 shows a diagram representing the sequence of steps we follow when pre-processing each tweet. Figure 3 displays the sequence of steps we follow when clustering tweets.

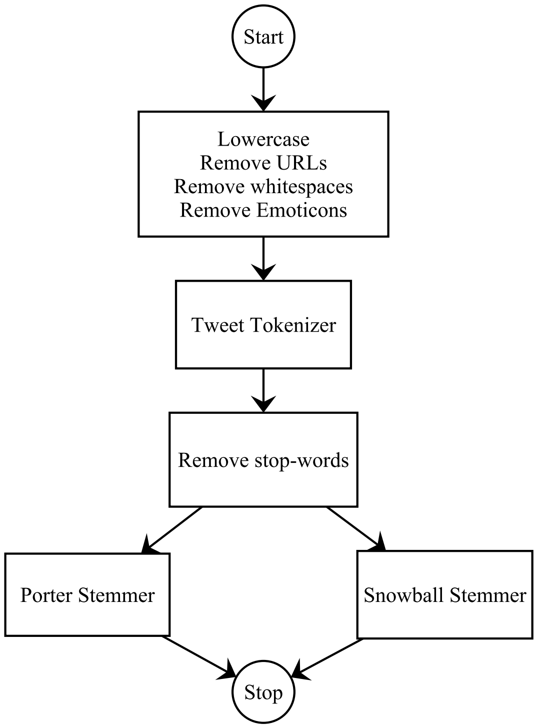

Fig. 2. Tweet Pre-Processing.

In preparation for trend discovery, we divided the corpus into *time-slices*. Then, we clustered the tweets within each time slice separately. The idea behind clustering is that tweets belonging to the same trend will cluster together. Hence, we can consider each cluster as a potential trend.

As indicated in Figure 3, the first step after separating the tweets into time-slices, and pre-processing them, is feeding them into a `TfidfVectorizer` from *scikit-learn*, a freely-available machine learning library for Python [54]. The `TfidfVectorizer` converts each tweet into a matrix of TF-IDF features. In principle, the only two features we took into account were the bigrams and trigrams available in the tweets for each time-slice. We avoided unigrams, because they are too limited to characterise a trend [55].

Although we extracted all the bigrams and trigrams available in the corpus, we discarded those which occur in only one tweet. This allowed us to reduce the "vocabulary" to a manageable size. Indeed, removing the bigrams and trigrams which occur in only one tweet within the first 10,000 tweets in the corpus, reduced the total from 186,850 to 43,885. This reduction removed almost 75% of the total vocabulary.

We limited the time slices to 10 minutes: the first slice covers from 14:24, which is the start of the retrieval, to 14:34; the second one covers from 14:34 to 14:44; and so on. We could have chosen larger slices—for example, Palomino *et al.* used 30-minute slices in the past [56]. However, the vocabulary can grow so much, and so fast, that larger slices inevitably increase the complexity and slow down the execution.

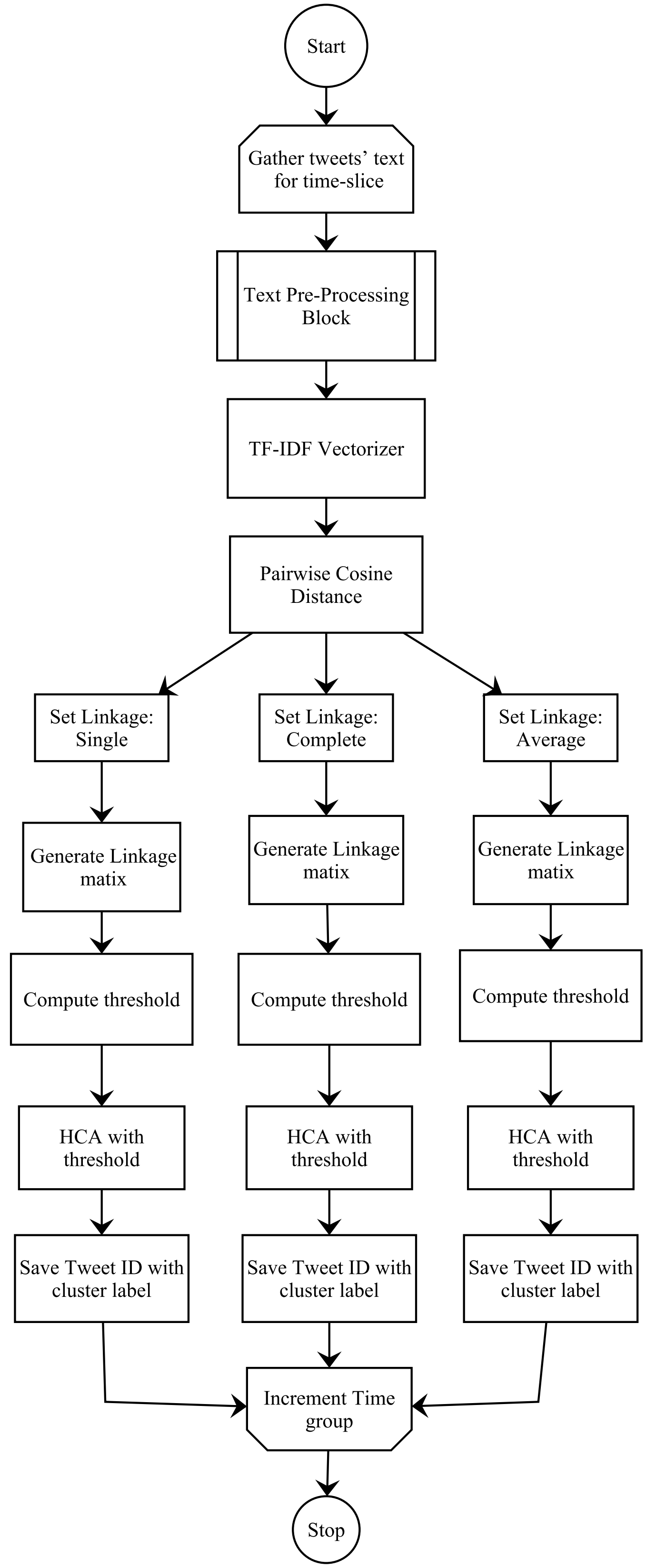

Fig. 3. Clustering.

Once the `TfidfVectorizer` for each tweet in the time-slice is ready, we use the cosine similarity metric to compute the distance between each pair of tweets. This yields a square matrix whose order is the number of tweets in the time-slice. Such a matrix records the pairwise-similarity and distance between any pair of tweets. The agglomerative hierarchical clustering algorithm, which we use to cluster the tweets, begins with this matrix. Initially, all the tweets are considered to be in their own cluster. Then, the algorithm merges, iteratively, the two closest clusters into a new one.

Although we tested the three variants of linkage—namely, single, complete and average—, we chose the complete-link variant in the end, because it performed better than the others, which agrees with former studies dedicated to clustering [57]. We explain below how we validated our clusters.

A. Trend Validation

We validated the performance of our trend discovery algorithm by means of retrieving a second corpus, which we refer to as the *supporting* corpus. Such a corpus helped us to corroborate if the clusters we obtained were indeed trends.

To form the supporting corpus, we launched a second retrieval of tweets at the same time we carried out the first one. However, the second retrieval captured, exclusively, tweets published by verified Twitter accounts owned by popular news and media sources in English language. We chose a total of 96 of such accounts, and they are all listed in Table II. On average, each of these accounts has more than 5 million followers, and the first and last tweet published by them were retrieved at 14:24:34 and 18:48:27 on 22 April 2020, respectively. The total number of tweets retrieved was 1,771. Figure 4 depicts the number of tweets we retrieved every 30 minutes.

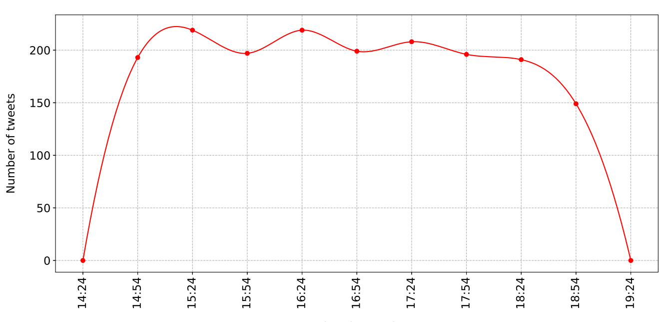

Fig. 4. Supporting corpus: Distribution of tweets over the retrieval.

Given the nature of the accounts in Table II, we can expect the tweets composing the supporting corpus to be mostly breaking news and headlines published at the same time of the press briefing we monitored. If one of our clusters matched the breaking news and headlines, we concluded such a cluster was an actual trend. In other words, if the information mentioned in one of our clusters was identified and tweeted about by a well-regarded news source, such as *BBC News*, we assumed we accurately detected a trend.

The tweets in the supporting corpus were also divided into the same time-slices, and were also pre-processed by changing them to lowercase and removing the stop words. Then, we extracted, for each cluster in the experimental corpus, the three most frequent bigrams and trigrams, and we compared them with the bigrams and trigrams derived from the tweets in the supporting corpus. If a match was found, the corresponding cluster was confirmed to be a trend.

Figure 5 compares the total number of clusters formed in the corpus with the number of actual trends detected every 10 minutes. In total, 54 of the 126 clusters discovered were confirmed to be trends. An example of one of the trends we identified is displayed in Table III. Note that Table III also indicates the number of tweets in the trend, and the most characteristic bigrams and trigrams found in the trend.

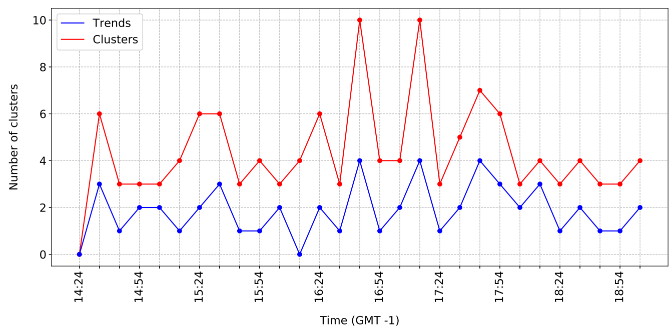

Fig. 5. Clusters vs trends.

B. Net Sentiment Rate

To assess the impact of sentiment on trend discovery, we calculated the *net sentiment rate* (NSR) for all our clusters. The NSR is a metric to estimate the overall sentiment expressed towards particular topics on social networks [58]. It is defined as the difference between the number of positive conversations—positive tweets in our case—and the number of negative conversations—negative tweets—divided by the total number of conversations—total number of tweets:

$$NSR = \frac{Positive\ tweets - Negative\ tweets}{Total\ number\ of\ tweets}$$

As indicated above, we selected SentiStrength and TextBlob to determine the sentiment expressed on the tweets constituting the clusters across all the time-slices. We selected two tools, because we are aware of the lack of consensus in sentiment analysis tools [42], and we wanted to prevent any biases derived from choosing a single one.

Figure 6 displays the number of positive, negative and neutral tweets in the experimental corpus, according to SentiStrength and TextBlob. Evidently, the two tools disagree overall: SentiStrength encounters more negative tweets than positive ones, whereas TextBlob encounters exactly the opposite. However, when calculating the NSR for the trends discovered, both tools recognise a similar pattern among trends. We elaborate on this in the following section.

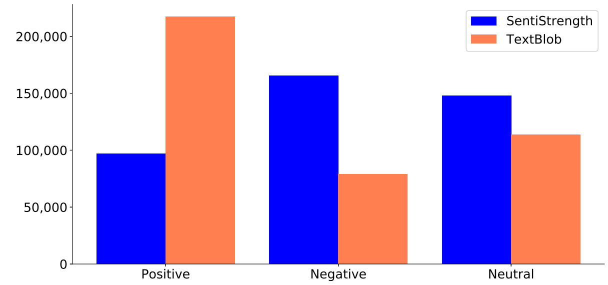

Fig. 6. Polarity over the entire experimental corpus.

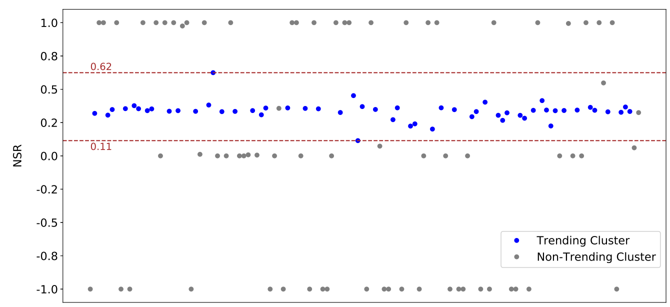

Fig. 8. NSR for trending clusters (according to TextBlob).

V. Results

Figure 7 shows the NSR for all the clusters available in the experimental corpus according to SentiStrength. Note that the actual trends are shown in blue colour, whereas the clusters which were not identified as trends are shown in grey. Also, note that the NSR is a number between -1 (totally negative) and 1 (totally positive). Zero represents a neutral NSR.

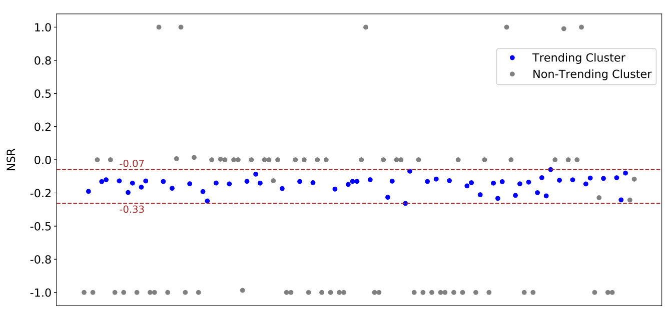

Fig. 7. NSR for trending clusters (according to SentiStrength).

As shown in Figure 7, the NSR for all the trends, according to SentiStrength, is consistently between -0.33 and -0.07, which means the actual trends are barely negative. This seems to contradict our initial assumption that extreme polarities, as opposed to neutral ones, are required to generate a trend. Moreover, Figure 7 also proves that the clusters which have extreme polarities—those whose NSR is either -1 or 1—were not recognised as trends.

A similar behaviour is found after determining the NSR with a different tool. Figure 8 shows the NSR for all the clusters in the experimental corpus according to TextBlob. Once again the actual trends are shown in blue colour, whereas the clusters which were not identified as trends are shown in grey. Recall that TextBlob considers the corpus more positive than SentiStrength—see Figure 6. Thus, the trends in Figure 8 are positive; yet, they are not significantly positive. Indeed, all the trends but one have an NSR lower than 0.5. A clear correlation between confirmed trends and strength of sentiment cannot derived from our experiments.

The NSR plots displayed in Figure 7 and Figure 8 suggest that clusters which are predominantly neutral are the most trending. This is in direct contradiction with our initial assumption that extreme polarities are required to generate a trend. Surprisingly, we have also found that all the trends fall within a narrow NSR range, regardless of the tool used to determine the NSR.

VI. Conclusions

Further research is necessary to study the relationship between trends and strength of sentiment in Twitter. Our work has demonstrated that highly polarised tweets are not the only ones capable of forming a trend: neutral tweets can become trends too, under certain circumstances.

It seems feasible that certain domains, such as the domain of our experimental corpus, are more likely to favour the presence of neutral trends. While controversial comments about COVID-19 have been published on Twitter—for example, regarding the Swedish strategy to avoid lockdown or the Lancet's much-publicised paper on the use of a chloroquine-based treatment [59]—, our experimental corpus is largely factual and descriptive. Hence, we cannot expect a considerable presence of disruptive and polarising tweets in our corpus, which could, in turn, drive the emergence of trends.

While a classifier could be trained on our data to predict whether a future cluster can become a trend, our research has evidenced the need for further research on the nature of trends in Twitter. The existing literature does not document in sufficient depth the features of the sentiment expressed in trends, their emergence and evolution over time.

Acknowledgements

Marco Palomino acknowledges the funding provided by the Interreg 2 Seas Mers Zeeën AGE'IN project (2S05-014).

References

[1] J. Min, M. Ziegler, H. J. J. Chang, and T.-M. C. C. Jai, "The Effects of Celebrity-Brand Congruence and Publicity on Consumer Attitudes, Buying Intention, and Loyalty," in *Proceedings of the International Textile and Apparel Association Annual Conference*, vol. 74, no. 1. Iowa State University Digital Press, 2017.

[2] J. Berger, A. T. Sorensen, and S. J. Rasmussen, "Positive Effects of Negative Publicity: When Negative Reviews Increase Sales," *Marketing Science*, vol. 29, no. 5, pp. 815–827, 2010.

TABLE II
TWITTER NEWS AND MEDIA ACCOUNTS MONITORED

@ABC	@AJEnglish	@AP	@AssociatedPress	@AxiosWorld	@BBCNews
@BBCSport	@BIUK	@BleacherReport	@BreibartNews	@BuzzFeed	@CBCNews
@CBSNews	@Conservatives	@DailyMailUK	@DailyMailUS	@DailyMirror	@EW
@FOXSports	@FT	@FinancialReview	@FinancialTimes	@FortuneMagazine	@FourFourTwo
@FourFourTwoUSA	@FoxNews	@GovUK	@GuardianAus	@HuffPost	@Independent
@IrishTimes	@LibDems	@MSNBC	@MTVNews	@MailOnline	@MetroUK
@NBCNews	@NFLNewsdesk	@NHSUK	@NRO	@NYMag	@NatGeo
@News24	@Newsweek	@Plaid_Cymru	@Polygon	@Recode	@Reuters
@ReutersUK	@TIME	@TechCrunch	@Telegraph	@TelegraphNews	@TheEconomist
@TheGreenParty	@TheNextWeb	@TheSNP	@UKLabour	@USATODAY	@UUPOnline
@WIRED	@WSJ	@WashTimes	@WiredUK	@abcnews	@arstechnica
@axios	@business	@businessinsider	@engadget	@espn	@financialpost
@foxnewsalert	@ftlive	@globeandmail	@googlenews	@guardian	@ladbible
@mashable	@mnt	@newscientist	@newscomauHQ	@nytimes	@politico
@rte	@sinnfeinireland	@sportbible	@talkSPORT	@techradar	@the_hindu
@thehill	@timesofindia	@verge	@vicenews	@washingtonpost	

TABLE III
EXAMPLES OF TRENDING TOPICS AND THEIR ASSOCIATED TWEETS

Cluster 1:	219 tweets in total
Terms:	areas coronavirus; coronavirus covid19; coronavirus covid19 covid-19; covid 19;
Terms:	covid19 covid-19;
Comment:	All the tweets in this cluster refer to coronavirus in various geographical areas: London, Glasgow, Aberdeen, California, China, Detroit, LA County, Sumatra, Vietnam, Nova Scotia...
Tweets:	London bus drivers' action over safety stops front door boarding "Drivers are still...
	Major new coronavirus testing lab opens in Glasgow https://t.co/IJBcu2GEi8 #coronavirus
	Aberdeen archaeologist faces freezing temperatures and prowling wolves during lockdown...
	REGIONAL BREAKDOWN: Further 43 people diagnosed with Covid-19 in north and north-east...
	@CNN Airplanes were arriving in #California from #Wuhan #China during that time...
	New: A nurse in Detroit was fired after speaking about working conditions inside...
	Testing shows hundreds of thousands in LA County may have been infected with ...
	Update: COVID-19: Scientists in South Sumatra claim glucose-based snack 'cures' ...
	#Vietnam to ease lockdown Thursday in most areas https://t.co/pCqtNLbs0m ...
	Coronavirus, Antibody Tests, Nova Scotia: Your Monday Briefing READ MORE: ...
	A Kashmiri boy looking through windowpane as he covers his face with a polythene...
	What cities near me have been affected by coronavirus? via @aajtak #coronavirus...
	#Coronavirus #COVID19 #Greece locks down migrant hotel after virus...
	Here We Have A Florida Woman Explaining Why Miami Is To Blame For The Coronavirus!
	Why getting the U.S. back to normal in the next couple months is a 'fantasy'...
	CAMBRIDGE TEMPORARY #HOMELESS #SHELTER OPENS DAY LATE AFTER TESTING MANDATE...
	Leeds City Council provide #Coronavirus #Covid19 videos available in #Polish...
	#KENYA- The coronavirus, #COVID19 has led to an increase in online food purchases...
	...

[3] X. Fang and J. Zhan, "Sentiment Analysis Using Product Review Data," *Journal of Big Data*, vol. 2, no. 5, pp. 2196–1115, 2015.

[4] M. Hu and B. Liu, "Mining and Summarizing Customer Reviews," in *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY: Association for Computing Machinery, 2004, p. 168–177. [Online]. Available: https://doi.org/10.1145/1014052.1014073

[5] M. Mertiya and A. Singh, "Combining Naive Bayes and Adjective Analysis for Sentiment Detection on Twitter," in *Proceedings of the 2016 International Conference on Inventive Computation Technologies (ICICT)*, vol. 2. IEEE, 2016, pp. 1–6.

[6] K. Mishra and A. Mishra, "Building Brand Engagement through Twitter Marketing Strategies," *Journal of Digital & Social Media Marketing*, vol. 4, no. 4, pp. 380–390, 2017.

[7] D. King, D. Ramirez-Cano, F. Greaves, I. Vlaev, S. Beales, and A. Darzi, "Twitter and the Health Reforms in the English National Health Service," *Health Policy*, vol. 110, no. 2-3, pp. 291–297, 2013.

[8] D. Murthy, *Twitter: Social Communication in the Twitter Age*. Polity Press Cambridge, UK, 2018.

[9] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," in *International Conference on Language Resources and Evaluation*, vol. 10, no. 2010, 2010, pp. 1320–1326.

[10] Y. Lin, "10 Twitter Statistics Every Marketer Should Know in 2020 [Infographic]," *Oberlo Blog*, 2019.

[11] M. Hasan, M. A. Orgun, and R. Schwitter, "A Survey on Real-Time Event Detection from the Twitter Data Stream," *Journal of Information Science*, vol. 44, no. 4, pp. 443–463, 2018.

[12] M. Palomino and A. Murali, "#Brexit Vs. #StopBrexit: What is Trendier? An NLP Analysis," in *Proceedings of the 8th International Conference on Natural Language Processing*, Copenhagen, Denmark, 2019, pp. 19–34.

[13] Office for National Statistics, "Social Networking by Age Group, 2011 to 2017," 2017. [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/adhocs/007401socialnetworkingbyagegroup2011to2017

[14] D. Boyd, S. Golder, and G. Lotan, "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter," in *Proceedings of the*

43rd Hawaii International Conference on System Sciences (HICSS). Honolulu, HI: IEEE, Mar. 2010, pp. 1–10.

[15] M. J. Paul and M. Dredze, "You Are What You Tweet: Analyzing Twitter for Public Health," *Proceedings of the International Conference on Weblogs and Social Media*, vol. 20, pp. 265–272, 2011, Barcelona, Spain.

[16] P. Velardi, G. Stilo, A. E. Tozzi, and F. Gesualdo, "Twitter Mining for Fine-Grained Syndromic Surveillance," *Artificial Intelligence in Medicine*, vol. 61, no. 3, pp. 153–163, 2014.

[17] V. Lampos and N. Cristianini, "Tracking the Flu Pandemic by Monitoring the Social Web," in *Proceedings of the 2nd International Workshop on Cognitive Information Processing (CIP)*. IEEE, 2010, pp. 411–416.

[18] A. Culotta, "Towards Detecting Influenza Epidemics by Analyzing Twitter Messages," in *Proceedings of the 1st Workshop on Social Media Analytics*. Washington, DC: ACM, 2010, pp. 115–122.

[19] E. De Quincey and P. Kostkova, "Early Warning and Outbreak Detection Using Social Networking Websites: The Potential of Twitter," in *Proceedings of the International Conference on Electronic Healthcare*. Springer, 2009, pp. 21–24.

[20] D. Scanfeld, V. Scanfeld, and E. L. Larson, "Dissemination of Health Information through Social Networks: Twitter and Antibiotics," *American Journal of Infection Control*, vol. 38, no. 3, pp. 182–188, 2010.

[21] K. Wilson and J. S. Brownstein, "Early Detection of Disease Outbreaks Using the Internet," *CMAJ*, vol. 180, no. 8, pp. 829–831, 2009.

[22] Twitter, Inc., "Twitter Trends FAQs," 2020. [Online]. Available: https://help.twitter.com/en/using-twitter/twitter-trending-faqs

[23] J. Benhardus and J. Kalita, "Streaming Trend Detection in Twitter," *International Journal of Web Based Communities*, vol. 9, no. 1, pp. 122–139, 2013.

[24] B. O'Connor, M. Krieger, and D. Ahn, "TweetMotif: Exploratory Search and Topic Summarization for Twitter," in *Proceedings of the 4th International AAAI Conference on Web and Social Media*, Washington, DC, 2010, pp. 384–385.

[25] H. Becker, M. Naaman, and L. Gravano, "Beyond Trending Topics: Real-World Event Identification on Twitter," *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*, vol. 11, no. 2011, pp. 438–441, 2011.

[26] S. Petrović, M. Osborne, and V. Lavrenko, "Using Paraphrases for Improving First Story Detection in News and Twitter," in *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, 2012, pp. 338–346.

[27] S. Petrovic, M. Osborne, R. McCreadie, C. Macdonald, I. Ounis, and L. Shrimpton, "Can Twitter Replace Newswire for Breaking News?" in *Proceedings of the 7th International AAAI Conference on Weblogs and Social Media (ICWSM)*, Boston, MA, 2013.

[28] A. Guille and C. Favre, "Event detection, tracking, and visualization in twitter: A mention-anomaly-based approach," *Social Network Analysis and Mining*, vol. 5, no. 1, p. 18, 2015.

[29] S. Liang and M. de Rijke, "Burst-Aware Data Fusion For Microblog Search," *Information Processing & Management*, vol. 51, no. 2, pp. 89–113, 2015.

[30] G. Salton and D. Harman, *Information Retrieval*. John Wiley and Sons Ltd., 2003.

[31] A. Rajaraman and J. D. Ullman, *Mining of Massive Datasets*. Cambridge University Press, 2011.

[32] D. A. Shamma, L. Kennedy, and E. F. Churchill, "Peaks and Persistence: Modeling the Shape of Microblog Conversations," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. ACM, 2011, pp. 355–358.

[33] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.

[34] R. Feldman, "Techniques and Applications for Sentiment Analysis," *Communications of the ACM*, vol. 56, no. 4, pp. 82–89, 2013.

[35] E. Cambria, "Affective Computing and Sentiment Analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.

[36] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, "Affective Computing and Sentiment Analysis," in *A Practical Guide to Sentiment Analysis*. Springer, 2017, pp. 1–10.

[37] B. Schuller, A. Batliner, S. Steidl, and D. Seppi, "Recognising Realistic Emotions and Affect in Speech: State of the Art and Lessons Learnt from the First Challenge," *Speech Communication*, vol. 53, no. 9-10, pp. 1062–1087, 2011.

[38] B. Liu, "Sentiment Analysis and Opinion Mining," *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[39] F. N. Ribeiro, M. Araújo, P. Gonçalves, M. A. Gonçalves, and F. Benevenuto, "Sentibench—A Benchmark Comparison of State-of-the-Practice Sentiment Analysis Methods," *EPJ Data Science*, vol. 5, no. 1, pp. 1–29, 2016.

[40] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, "Sentiment Strength Detection in Short Informal Text," *Journal of the American Society for Information Science and Technology*, vol. 61, no. 12, pp. 2544–2558, 2010.

[41] S. Loria, "TextBlob Documentation," *Release 0.15*, vol. 2, 2018.

[42] A. Connelly, V. Kuri, and M. Palomino, "Lack of Consensus among Sentiment Analysis Tools: A Suitability Study for SME Firms," in *Proceedings of the 8th Language & Technology Conference*, Poznan, Poland, 2017, pp. 54–58.

[43] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment Strength Detection for the Social Web," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 1, pp. 163–173, 2012.

[44] T. De Smedt and W. Daelemans, "Pattern for python," *Journal of Machine Learning Research*, vol. 13, no. 66, pp. 2063–2067, 2012.

[45] J. Perkins, *Python 3 Text Processing with NLTK 3 Cookbook*. Packt Publishing Ltd., 2014.

[46] NLTK 3.3. (2018) NLTK.Tokenize Package. [Online]. Available: http://www.nltk.org/api/nltk.tokenize.html#module-nltk.tokenize.casual

[47] GOV.UK, "Slides, Datasets and Transcripts to Accompany Coronavirus Press Conferences," 2020. [Online]. Available: https://www.gov.uk/government/collections/slides-and-datasets-to-accompany-coronavirus-press-conferences

[48] F. Walsh, "Coronavirus: First Patients Injected in UK Vaccine Trial," 2020, BBC News Services. [Online]. Available: https://www.bbc.co.uk/news/health-52394485

[49] J. Roesslein, "Tweepy Documentation," 2020. [Online]. Available: http://docs.tweepy.org/en/v3.5.0/

[50] F. Morstatter, J. Pfeffer, H. Liu, and K. M. Carley, "Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose," in *Proceedings of the 7th International AAAI Conference on Weblogs and Social Media*. Cambridge, MA: The AAAI Press, 2013, pp. 400–408.

[51] Twitter, Inc., "Filter Realtime Tweets," 2020. [Online]. Available: https://developer.twitter.com/en/docs/tweets/filter-realtime/overview

[52] ——, "Tweet Object," 2020. [Online]. Available: https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/tweet-object

[53] E. Bisong, "Google Colaboratory," in *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Springer, 2019, pp. 59–64.

[54] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[55] L. M. Aiello, G. Petkos, C. Martin, D. Corney, S. Papadopoulos, R. Skraba, A. Göker, I. Kompatsiaris, and A. Jaimes, "Sensing Trending Topics in Twitter," *IEEE Transactions on Multimedia*, vol. 15, no. 6, pp. 1268–1282, 2013.

[56] M. A. Palomino, Q. Ribac, and G. L. Masala, "The Nature of Twitter Trending Topics - Analysing Intrinsic Factors Associated with the Twitter Ecosystem," in *Proceedings of the 17th International Conference on New Trends in Intelligent Software Methodologies, Tools and Techniques*, Granada, Spain, 2018, pp. 1004–1017. [Online]. Available: https://doi.org/10.3233/978-1-61499-900-3-1004

[57] E. M. Voorhees, "The effectiveness and efficiency of agglomerative hierarchic clustering in document retrieval," Cornell University, Tech. Rep. TE 85-705, 1985.

[58] M. Palomino, T. Taylor, A. Göker, J. Isaacs, and S. Warber, "The online dissemination of nature–health concepts: Lessons from sentiment analysis of social media relating to "nature-deficit disorder"," *International Journal of Environmental Research and Public Health*, vol. 13, no. 1, p. 142, 2016.

[59] Y. Guzman-Prado, "Retraction of Studies on Potential Drug Therapies for Covid-19: A Call for Reliability and Scientific Integrity," *American Journal of Cardiology*, 2020.