Using Orders of Magnitude Reasoning to Aggregate and Compare News Reporting Sentiment

Jennifer Nguyen[1] and **Albert Armisen**[2] and **Nuria Agell**[3] and **Àngel Saz**[4]

Abstract. This paper focuses on analyzing the underlying sentiment of news articles, taken to be factual rather than comprised of opinions. The sentiment of each article towards a specific theme can be expressed in orders of magnitude terms and aggregated into a centralized sentiment which can be trended. This allows the interpretation of sentiments without conversion to numerical values. The methodology, as defined, maintains the range of sentiment articulated in each news article per day. In addition, a measure of consensus is defined for each day as the degree to which the articles published agree in terms of the sentiment presented. A real case example is presented for a controversial event in recent history with the analysis of 82,054 articles over a three day period. The analysis compares the internal consensus per day of different countries.

1 Introduction

Exploiting sentiment analysis for stock predictions [3], elections [8], sentiment tracking [6], has become common. Different sentiment techniques have been developed to identify sentiment in data such as blogs, online reviews, and microblogs [20]. Methodologies have been applied to trending sentiments towards particular topics such as election candidates, monitoring customer sentiment towards a product or business, aggregating customer reviews and so on. Many of these methods begin by assessing sentiment of written text on a n-point scale or as positive, negative or neutral labels [16]. However, this initial assessment can loose the original sentiment contained within the text if the positive and negative values assigned to each word in a text are summed together for an overall score. In addition, the sentiment can be further removed when the sentiment of groups of texts are aggregated. As sentiment can be both positive and negative in a group of texts, the average sentiment may appear neutral. Furthermore, the extent to which an individual text may be positive or negative is no longer evident in the mean. To this end, the methodology presented in this paper proposes to represent sentiment in order of magnitude terms to capture the range of positiveness and negativeness in an individual written text.

As the consumption of fake news becomes more and more of a concern, the professional role of media is at the forefront. Sensationalism contradicts the role of the news to report accurate information about events [7]. Journalists control what is explained and how the story is framed [11, 1]. This framing shapes how an audience discusses a story. However, its influence can diminish with an audience's awareness of issues or events in their communities or the world [4]. In addition, research has demonstrated that newspapers have the ability to alter the perception of the importance and dominant opinion a community associates with specific issues [15]. As such, newspapers may be able to indirectly change public opinion [15].

Public opinion can result in different actions. Therefore, there has been interest in studying investor sentiment [3], voter sentiment [8, 21], sentiment tracking [6]. These sentiments have been studied to predict stock forecasts [3], election outcomes [8], voter social media behavior [21], and detecting events related to sentiment change [6]. Some methods applied a mean sentiment score [3, 21], others analyzed the frequency with which texts scored a specific sentiment [8, 6], others a moving average [6]. For some papers, the overall sentiment scores were defined for each tweet as the sum of the positive and negative sentiments assigned [21, 6]. Another paper, traced the divergence of opinions over time [2]. In order to study public sentiment, each of these studies focused on a particular theme such as political candidates, specific stocks or event. Lastly, these papers analyzed sentiment from Twitter posts rather than news articles, a platform where expressions of opinion are more accessible and possibly more obvious.

In this paper, we propose a methodology which highlights central sentiment and degree of consensus among a group of news articles. The sentiment of each article is expressed in order of magnitude terms. As centralized sentiment is computed to reflect country specific opinion and a comparison is made among the positions taken by each country. The approach is novel, to our knowledge, in considering article sentiment as intervals rather than separately, or as an average or substraction of positive and negative sentiment preserving the original opinion. In addition, sentiment is described in orders of magnitude terms to better reflect the natural manner in which humans discuss articles.

The rest of the paper is organized as follow. First, the preliminaries underlying the methodology are presented in Section 2. Next, the proposed methodology is introduced in Section 3. It is followed by a real case example and a discussion of the results from its implementation in Section 4. Lastly, the conclusion and future work are presented in Section 5.

2 Preliminaries

In this section, a summary of basic concepts related to Orders of Magnitude Term Sets (OMTS) which will be referenced in the methodology are presented. These concepts are considered following absolute orders of magnitude reasoning model introduced by Travé-Massuyès [19, 18].

[1] Ramon Llull University, Spain, email: jenniferthanhvan.nguyen@esade.edu
[2] Vic University - Central University of Catalonia, Spain, email: albert.armisen@uvic.cat
[3] Ramon Llull University, Spain, email: nuria.agell@esade.edu
[4] Ramon Llull University, Spain, email: angel.saz@esade.edu

Let $\mathcal{S}$ denote a finite totally ordered set of linguistic terms, $\mathcal{S} = \{a_1, \ldots, a_n\}$ with $a_1 < \cdots < a_n$. In this article, a OMTS is defined as a set $\{x \in \mathcal{S} \mid a_i \leq x \leq a_j\}$ that is denoted as $[a_i, a_j]$ if $i < j$ or $\{a_i\}$ if $j = i$. Note that the same concept is denoted as fuzzy linguistic term set in [17]. Then, $\mathcal{H}_\mathcal{S}$ is defined as the set of all possible OMTS over $\mathcal{S}$ including the empty OMTS, $\{\emptyset\}$, such that $\mathcal{H}^*_\mathcal{S} = \mathcal{H}_\mathcal{S} - \{\emptyset\}$, according to the extension of fuzzy linguistic terms set considered in Montserrat et al. [13].

The set $\mathcal{H}_\mathcal{S}$ is extended to $\overline{\mathcal{H}_\mathcal{S}}$, to include the concepts of *positive intersection of linguistic terms*, *negative intersection or gap of linguistic terms* and *zero linguistic terms*. *Positive intersection OMTS* come from two OMTS with some linguistic terms in common, *zero OMTS* are the result of the intersection of two consecutive OMTS, while *negative intersecion OMTS* are the result of two OMTS with no common or consecutive linguistic terms.

In addition, the extended connected union and extended intersection operators are considered in this context as explained in [13] in the case of fuzzy linguistic terms sets.

1. The *extended intersection* of H_1 and H_2, $H_1 \sqcap H_2$, is the largest element in $\overline{\mathcal{H}_\mathcal{S}}$ that is contained in H_1 and H_2.
2. The *extended connected union* of H_1 and H_2, $H_1 \sqcup H_2$, is the smallest element in $\overline{\mathcal{H}_\mathcal{S}}$ that contains H_1 and H_2.

The extended intersection and extended connected union can be used to compute the distance between two OMTS as defined in [13] for hesitant fuzzy linguistic terms sets. Given H_1 and $H_2 \in \overline{\mathcal{H}_\mathcal{S}}$, the *width* of H, $\mathcal{W}(H)$, is defined as the number of linguistic terms contained in H, or cardinality, $card(H)$, if $H \in \mathcal{H}_\mathcal{S}$ or $-card(-H)$ if H is a negative OMTS. Then the distance between OMTS in $\overline{\mathcal{H}_\mathcal{S}}$ is computed between H_1 and H_2, as:

$$D(H_1, H_2) := \mathcal{W}(H_1 \sqcup H_2) - \mathcal{W}(H_1 \sqcap H_2). \quad (1)$$

To obtain the central sentiment (or centroid) of a set of articles about a specific theme λ , the distance D is applied as follows:

Definition 1 *([13]) Let λ be a theme, G a set of r articles and $H_1, \ldots, H_r$ the OMTS expressed by the articles in G with respect to the theme, λ. Then, the centroid of the set is:*

$$C_o = \arg \min_{H \in \mathcal{H}^*_\mathcal{S}} \sum_{i=1}^{r} D(H, H_i). \quad (2)$$

The centroid is a central measure for ordinal scales with hesitancy. In addition to the centroid, we consider the consensus degree proposed by [12] to quantify the sentiment agreement among a set of articles.

Definition 2 *([14]) Let G be a set of r articles of a theme λ, and $H_1, \ldots, H_r$ be their respective sentiments in OMTS. Let C_o be the central sentiment of the set. Then, the degree of consensus of G on λ is defined as:*

$$\delta_\lambda(G) = 1 - \frac{\sum_{i=1}^{r} D(C_o, H_i)}{r \cdot (n-1)}. \quad (3)$$

Note that $0 \leq \delta_\lambda(G) \leq 1$ as $r \cdot (n-1)$ is an upper bound of the addition of distances between the centroid and sentiment expressed as OMTS [14].

3 The proposed approach to detecting contrasting sentiment

In this section, we present the formal framework to determine the centroid and consensus among the sentiments of articles. Generally, articles published on the same day about the same theme do not have to reflect the same sentiments. However when it does, it could indicate that the sources of the articles are motivated in the same direction. Further analysis could be performed by evaluating the consensus of articles from neighboring or allied countries. Likewise, articles from a specific country could be trended by representing the aggregate articles for each day in terms of their centroid. A spike in the any direction different from the trend can draw attention.

The methodology requires as input a set of articles previously identified with positive and negative sentiment, and themes. The process of identifying the sentiment and theme are considered out of scope of this methodology as we are focused on identifying the centroid and measuring the consensus. The process has four steps: 1) Select theme for analysis and the corresponding articles, 2) Represent sentiment in linguistic terms, 3) Identify the centroid, and 4) Measure the consensus as shown in Figure 1.

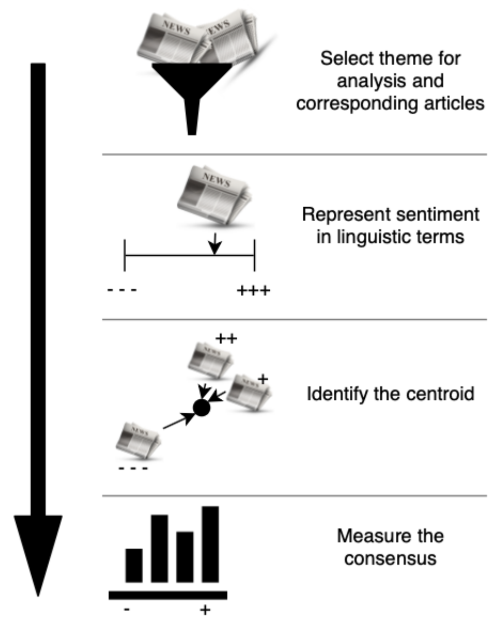

Figure 1. Framework to determine the centroid and consensus among article sentiment

1. *Select theme for analysis and the corresponding articles:* In this step, a theme is selected in order to focus the analysis. The data set is filtered for only those articles which reference a particular theme regardless of the degree to which a theme is mentioned.

2. *Represent sentiment in linguistic terms:* Next, for each article there are positive and negative sentiments. If the article is associated with positive and negative sentiment for each word in the article, the percent of negative and percent of positive words needs to be computed. We will refer to these percentages as

positive and negative scores going forward.

Different from the methodologies previously discussed in Section 1, the methodology presented in this paper proposes to utilize intervals to represent the article sentiments. Intervals assist with distinguishing cases in which you have a polarization in sentiment. For example, an article with a positive score, +12, and negative score, -11, could be summarized by its average sentiment, 0.5. Similarly, an article with a positive score, +5, and negative score, -3.5, would be summarized by its average sentiment, 0.75. Both of these examples would appear neutral. However, an interval would highlight that the first article expressed extreme sentiment in both directions. Whereas, the second article communicated with lesser sentiment. In addition, each interval is converted to linguistic terms to better represent how humans would describe an article.

Given an article with negative and positive scores, A^- and A^+, respectively, the sentiment can be represented in linguistic terms as:

$$H_A = min\{H \in \mathcal{H}_\mathcal{S} / [A^-, A^+] \subset H\}, \tag{4}$$

when $A^- \neq 0$ and $A^+ \neq 0$. In the case that $A^- = 0$ or $A^+ = 0$, then H_A is $min\{H \in \mathcal{S}/A^- \in \mathcal{S}\}$ or $min\{H \in \mathcal{S}/\ A^+ \in \mathcal{S}\}$, respectively.

Example 3.1 *Let us consider a set of possible sentiments in linguistic terms:* $\mathcal{S} = \{$*very negative, negative, somewhat negative, somewhat positive, positive, very positive*$\}$ *where very negative(VN)* $= [-100, -10]$*, negative(N)* $= (-10, -5]$*, somewhat negative(SN)* $= (-5, 0]$*, somewhat positive(SP)* $= (0, 5]$*, positive(P)* $= (5, 10]$*, very positive(VP)* $= (10, 100]$ *from which an article's sentiment may be described. Given an article with positive score* $A^+ = 3$ *and negative score* $A^- = -8$*, the representation of the sentiment of the article in linguistic terms would be* [*N, SP*]*. Similarly, given an article with positive score* $A^+ = 12$ *and negative score* $A^- = 0$*, the representation of the sentiment of the article in linguistic terms would be* [*VP*].

3. *Identify the centroid:* Once all the pairs of positive and negative scores for each article in the set have been translated into linguistic terms, the centroid can be computed according to Equation 2 and distance D from Equation 1. This represents the central sentiment of the set of articles.

Example 3.2 *Let us consider* G *to be a set of 5 articles written about a theme* λ*. The sentiment of each article is expressed in OMTS over the set* $\mathcal{S}$ *from Example 3.1. If* H_1, H_2, H_3, H_4, H_5 *are the OMTS of the sentiment communicated in the 5 articles, then the centroid of the set of articles,* C_o*, can be identified as shown in Table 1.*

	H_1	H_2	H_3	H_4	H_5	$\boldsymbol{C_o}$
λ	$[N, SN]$	$\{N\}$	$[SP, P]$	$[VN, N]$	$[N, SP]$	$\boldsymbol{[N, SN]}$

Table 1. Centroid of the set of articles G related to theme λ.

4. *Measure the consensus:* To understand to what extent articles in the set share similar sentiment, we compute the distance of each one of them to the central sentiment and determine its consensus δ from Equation 3.

Example 3.3 *Continuing with Example 3.2, the distances,* D*, between the centroid* C_o *and the sentiment of each article are computed using Equation 1. The distances are shown in Table 2 along with their associated degrees of consensus.*

	D_1	D_2	D_3	D_4	D_5	$\sum_{i=1}^{5} D_i$	$\delta_\lambda(G)$
λ	0	1	4	2	2	9	0.45

Table 2. Consensus of sentiments for articles in set G related to theme λ

4 A real case example

We demonstrate the presented methodology works with a subset of data from GDELT [10], an open news platform. "GDELT monitors print, broadcast, and web news media in over 100 languages from across every country in the world..."[5]. Its archives are continuously updated every 15 minutes providing information on the people, locations, organizations, themes, sources, emotions, counts, quotes, images and events discussed in each article. GDELT has been used in previous research studies related to global news coverage of disasters [9], effects of political conflict [22], and predicting social unrest [5].

The GDELT data set is publicly available[6]. For each article, a list of all the themes found in the document are provided. There are 2589 possible themes from which a document can be labeled. The data set includes the tone or sentiment for each of the articles. The tone for an article is described in terms of six emotional dimensions: the average tone of the article, the positive score, negative score, percentage of words found in the tonal dictionary, percentage of active words, and percentage of self/group reference. The average tone is the positive score minus the negative score. The positive and negative scores are the percentage of all words in the article found to have positive and negative emotion, respectively.

To illustrate the viability of the methodology, we selected data before, during, and after a controversial news event. Specifically, we selected January 7 through January 9, 2020 during which time the crash of a Ukrainian airplane was being questioned. During this time frame, information for 589,815 articles were collected from GDELT's database. The articles were filtered for those labeled as "Conflict and Violence" to narrow the articles to those most related to the event. This reduced the data set to 82,054 articles. The positive and negative scores for each of the articles were selected to represent the tone. These were selected as they represented the range and quantity of each type of sentiment present in an article, making them a better descriptor than the average tone, as previously mentioned.

Next, we represented each article's tone in linguistic terms following Equation 4 and Example 3.1. Table 3, shows the translation of five articles from the dataset.

---

[5] https://www.gdeltproject.org
[6] http://data.gdeltproject.org/gdeltv2/masterfilelist.txt

Article	Country	Date	Positive Score	Negative Score	Linguistic Term
1	US	07/01/2020	0.72	6.46	[N, SP]
2	US	07/01/2020	6.25	3.12	[SN, P]
3	US	07/01/2020	1.29	7.41	[N, SP]
4	US	07/01/2020	0	4.61	[SP, SP]
5	US	07/01/2020	1.32	8.07	[N, SP]

Table 3. Representation of sentiment in linguistic terms

Then, we analyzed the consensus on the tone of the articles to each country level. Therefore, the centroid was computed as the central sentiment for each of the three days selected for all the articles published. At this level we were able to analyze changes in sentiment by a given country during the selected event period. Table 4 depicts the same linguistic term that is consistent across different days in US, but the level of consensus changes across the different days.

Country	Date	Linguistic Term	Consensus
US	07/01/2020	[N,SP]	0.894
US	08/01/2020	[N,SP]	0.900
US	09/01/2020	[N,SP]	0.890

Table 4. Linguistic Terms & Consensus of the US for the three different days

5 Conclusion and Future Research

The proposed methodology translates positive and negative sentiment scores into linguistic terms. These terms are expressed as elements of the lattice of OMTS. This allows on the one hand the computation of distances among article sentiments and on the other hand the identification of the central representation together with the consensus of the sentiment. By using these linguistic terms, the methodology enables more explainable results compared to the results obtained when using numerical values. Regarding the future research, we considered comparing the centroid and consensus results from our proposed methodology to those obtained by means of a quantitative approach.

A limitation of the current methodology is that the analysis of the positive and negative sentiment of an individual article can be misinterpreted when coming from an individual theme. This is due to the sentiment scores of an article being associated with the entire article. Therefore, they cannot be separated into the different themes discussed. As such, the degree of positiveness or negativeness attributed to the theme analyzed cannot be certain. In this direction, as a future work, we will consider the use of different aggregation functions or feature modeling to take into account the imbalance among the themes represented in the articles. Furthermore, we plan to extend this methodology to take into account the relative importance of each segment or paragraph when weighting the scores. A second direction for our future work is to track sentiment trends to observe how different types of events affect the news reported.

ACKNOWLEDGEMENTS

We would like to thank the referees for their comments, which helped improve this paper considerably. This project has received funding from the European Union's Horizon 2020 research and innovation programme (grant agreement #822654)

REFERENCES

[1] Deseree LP Cain-Arzu, 'Sensationalism in newspapers: A look at the reporter and amandala in belize 2010–2014', (2016).

[2] Nan Cao, Lu Lu, Yu-Ru Lin, Fei Wang, and Zhen Wen, 'Socialhelix: visual analysis of sentiment divergence in social media', *Journal of visualization*, **18**(2), 221–235, (2015).

[3] Francesco Corea, 'Can twitter proxy the investors' sentiment? the case for the technology sector', *Big Data Research*, **4**, 70–74, (2016).

[4] Lauren Ducat, Stuart Thomas, and Warwick Blood, 'Sensationalising sex offenders and sexual recidivism: Impact of the serious sex offender monitoring act 2005 on media reportage', *Australian Psychologist*, **44**(3), 156–165, (2009).

[5] Divyanshi Galla and James Burke, 'Predicting social unrest using gdelt', in *International Conference on Machine Learning and Data Mining in Pattern Recognition*, pp. 103–116. Springer, (2018).

[6] Anastasia Giachanou and Fabio Crestani, 'Tracking sentiment by time series analysis', in *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pp. 1037–1040, (2016).

[7] Paul Hendriks Vettehen, Koos Nuijten, and Johannes Beentjes, 'News in an age of competition: The case of sensationalism in dutch television news, 1995–2001', *Journal of Broadcasting & Electronic Media*, **49**(3), 282–295, (2005).

[8] Sri Hari Deep Kolagani, Arash Negahban, and Christine Witt, 'Identifying trending sentiments in the 2016 us presidential election: A case study of twitter analytics', *Issues in Information Systems*, **18**(2), 80–86, (2017).

[9] Haewoon Kwak and Jisun An, 'A first look at global news coverage of disasters by using the gdelt dataset', in *International Conference on Social Informatics*, pp. 300–308. Springer, (2014).

[10] Kalev Leetaru and Philip A Schrodt, 'Gdelt: Global data on events, location, and tone, 1979–2012', in *ISA annual convention*, volume 2, pp. 1–49. Citeseer, (2013).

[11] Maxwell E McCombs and Donald L Shaw, 'The agenda-setting function of mass media', *Public opinion quarterly*, **36**(2), 176–187, (1972).

[12] J. Montserrat-Adell, N. Agell, M. Sánchez, and F. J. Ruiz, *A Representative in Group Decision by Means of the Extended Set of Hesitant Fuzzy Linguistic Term Sets*, 56–67, Springer International Publishing, Cham, 2016.

[13] Jordi Montserrat-Adell, Núria Agell, Mónica Sánchez, Francesc Prats, and Francisco Javier Ruiz, 'Modeling group assessments by means of hesitant fuzzy linguistic term sets', *Journal of Applied Logic*, **23**, 40–50, (2017).

[14] Jordi Montserrat-Adell, Núria Agell, Mónica Sánchez, and Francisco Javier Ruiz, 'Consensus, dissension and precision in group decision making by means of an algebraic extension of hesitant fuzzy linguistic term sets', *Information Fusion*, **42**, 1–11, (2018).

[15] Diana C Mutz and Joe Soss, 'Reading public opinion: The influence of news coverage on perceptions of public sentiment', *Public Opinion Quarterly*, 431–451, (1997).

[16] Rudy Prabowo and Mike Thelwall, 'Sentiment analysis: A combined approach', *Journal of Informetrics*, **3**(2), 143–157, (2009).

[17] Rosa M Rodriguez, Luis Martinez, and Francisco Herrera, 'Hesitant fuzzy linguistic term sets for decision making', *IEEE Transactions on fuzzy systems*, **20**(1), 109–119, (2011).

[18] Louise Travé-Massuyès, Liliana Ironi, and Philippe Dague, 'Mathematical foundations of qualitative reasoning', *AI magazine*, **24**(4), 91–106, (2003).

[19] Louise Travé-Massuyès, Francesc Prats, Mónica Sánchez, and Núria Agell, 'Relative and absolute order-of-magnitude models unified', *Annals of Mathematics and Artificial Intelligence*, **45**(3-4), 323–341, (2005).

[20] G Vinodhini and RM Chandrasekaran, 'Sentiment analysis and opinion mining: a survey', *International Journal*, **2**(6), 282–292, (2012).

[21] Ussama Yaqub, Soon Ae Chun, Vijayalakshmi Atluri, and Jaideep Vaidya, 'Sentiment based analysis of tweets during the us presidential elections', in *Proceedings of the 18th annual international conference on digital government research*, pp. 1–10, (2017).

[22] James Edward Yonamine, 'A nuanced study of political conflict using the global datasets of events location and tone (gdelt) dataset', (2013).