Decoding the Voices of Dissent: A Computational Analysis of Twitter Discourse during the Mahsa Amini Protests in Iran

Digital Social Data Research Project

University of Trento

Alessandro Cortese

27 January 2022

1 Introduction and Research Question

On 16 September 2022, the death of 22-year-old Mahsa Amini, a Kurdish woman arrested by Tehran morality police for alleged violations of Iran's strict veiling laws, has sparked widespread protests and social unrest across the country [1]. The demonstrations, which have taken place in hundreds of cities in Iran, have been met with harsh government repression, resulting at the time of writing in at least 476 deaths and over 18 thousand arrests [2].

The ongoing rallies have been driven by a variety of social, political end economic issues. Iranian citizens are protesting not only Amini's killing, but also the regime's escalating repression of women, journalists, and minorities. Moreover, harsh sanctions and rampant inflation in the country have contributed to increase resentment towards the Iranian government [3].

Despite extensive Internet censorship and communication blackouts, social media, particularly Twitter, have been playing a crucial role in the organization and dissemination of information about the demonstrations. Activists and ordinary citizens have used the platform to share news, coordinate actions, and express their views about the protests [4].

The aim of this paper is to understand which topics are more represented in the Twitter conversation around these multifaceted protests in order to gather insight on the issues driving them. In particular, we would like to understand:

- What is the demographic profile of the users tweeting about the Iran protests?
- What are the most common topics expressed in tweets about the protests? How do they change over time?
- What is the sentiment of such tweets? Does it reflect major events and developments in the protest movement?

2 Related Work

Twitter has become a commonly used tool for organizing and communicating among protesters globally, with notable examples of its impact in the Iranian election protests of 2009-2010 [5], the Egyptian revolution of 2011 [6] and the Occupy Wall Street protests in 2011 [7, 8]. The frequent use of the platform in political movements has led to the term "Twitter revolutions" being coined [5].

Therefore, it is unsurprising that Twitter has been considered by some as a catalyst for movements' success, with some claiming that more and more social movements became successful precisely because of social media [9]. For example, according to Grossman (2009), with respect to the 2009 Iranian presidential elections, “there’s no question that it [Twitter] has emboldened the protesters, reinforced their conviction that they are not alone and engaged populations outside Iran in an emotional, immediate way that was never possible before” [10].

As a consequence, the use of Twitter data as a tool for analyzing complex social movements, such as protests, has been the subject of much research in recent years [11]. For example, Tan *et al.* (2011) demonstrated the link between the vitality of the Occupy Wall Street movement and the volume of the related tweets over time, identifying “buzz makers” and allowing the trend of the movement to be predicted using Twitter data [7]. Zhou *et al.* (2010) studied information resonance on Twitter during the 2009 Iranian protests, showing that even though the followers network structure has a key role in propagating information via retweets, the most important channel for the spread of information were the suggested hashtags in the “Trending topics” section and on the search bar.

This approach has a number of advantages, one of them being the wide availability of data, which enables researchers to easily collect large volumes of information to be analyzed [12]. Additionally, the real-time nature of Twitter data allows scholars to obtain quick snapshots of the current situation, facilitating the identification of patterns and trends within social movements.

However, the use of Twitter as a means of studying social phenomena is not devoid of limitations. Firstly, Twitter data might not be representative of the broader population of interest or the movement as a whole. As a matter of fact, only a percentage of people use Twitter, and among those, only a minority uses the platform to tweet about specific movements [13]. Additionally, Twitter data is often biased towards certain groups and individuals, and may not capture the perspectives of marginalized or underrepresented groups. Finally, another concern comes from censorship, manipulation and fake news spreading on the platform, which can all skew the analysis.

3 Methodology

3.1 Dataset

The analysis is based on a set of 553’575 tweets whose publication date spans from 7 December to 30 December 2022. The data was collected using the Twitter API, focusing on tweets in English containing at least one of the following hashtags: #iran, #mahsaamini, #iranrevolution, #iranrevolution2022, #iranprotests, #iranprotests2022. Due to API limitations, which allow only for a maximum retrieval of 18000 tweets per request, tweets were separately collected and saved in different moments in time and then merged into a unique dataset.

The content of the tweets was then cleaned by applying case folding to lowercase, removing then digits and punctuation. Hashtags and mentions were also eliminated, along with stopwords. Links and other special symbols were removed. Moreover, a copy of this cleaned text was also subject to a stemming procedure, thus keeping only the root of each word.

3.2 Analysis

3.2.1 Exploratory Data Analysis

After the data collection procedure, some Exploratory Data Analysis has been carried out in order to perform an initial investigation and get a high-level overview of the data. Figure 1 represents a line chart of the daily tweet frequency over the considered time span.

It is possible to notice a series of spikes, indicating a large volume of tweets in some specific periods. The highest ones, surpassing 25000 daily tweets, revolve around 8-9, 12 and 20 December.

With respect to the first and by far the largest peak, it is worth noting that, on 8 December 2022, Iran carried out the first state-sponsored execution of the protests, hanging 23-year-old Mohsen Shekari as a punishment for obstructing a road and injuring a police officer during a demonstration

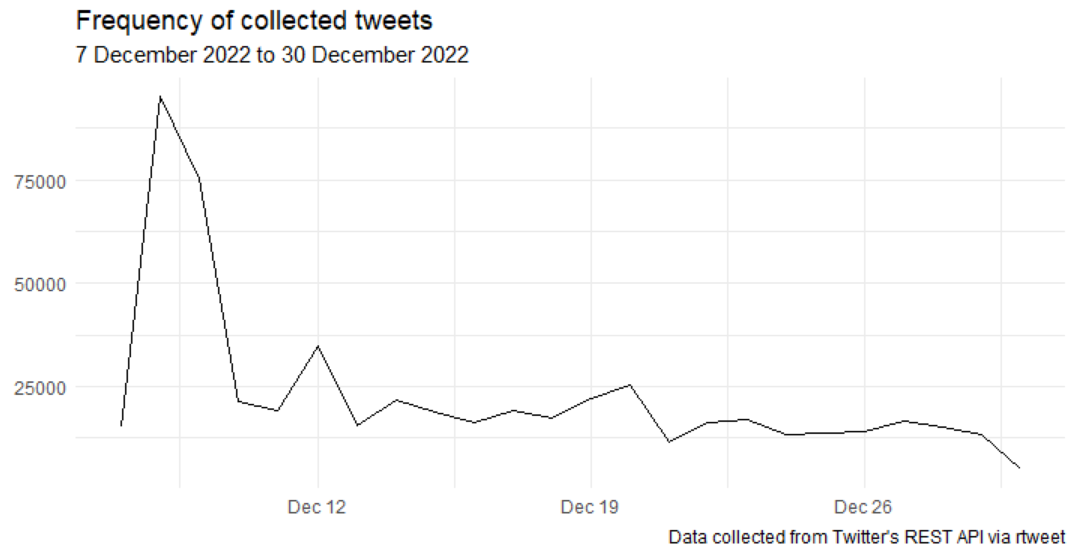
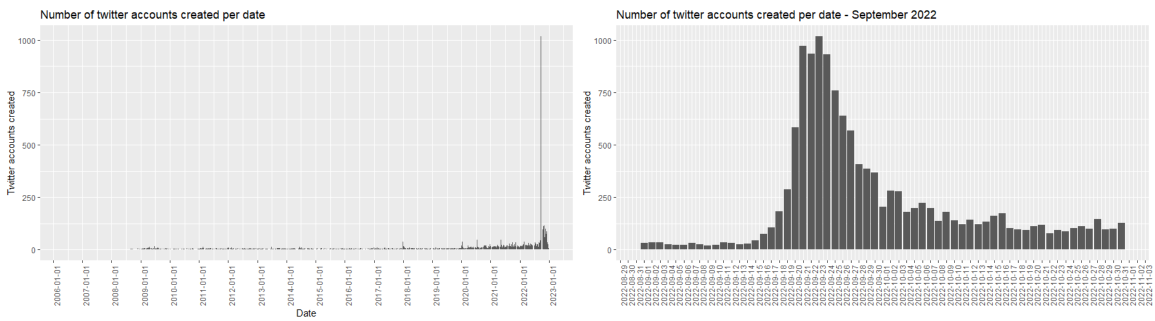


Figure 1: Daily frequency of collected tweets from 7 to 30 December 2022

in Tehran [14]. As a matter of fact, the hashtag #mohsenshekari appears to be the most frequent hashtag in the dataset, excluding those used for the scraping, appearing in total 100’453 times, 79’643 of which in tweets created between 8 and 9 December.

On 12 December, the most frequent hashtag, appearing 14768 times over the day, was #majidrezarahnavard, which is the name of the second protester executed by the Iranian government on that day [15].

The most common hashtag on 20 December was instead #mohammadmehdikarami, another protester who was actually executed on 7 January 2023 [16].

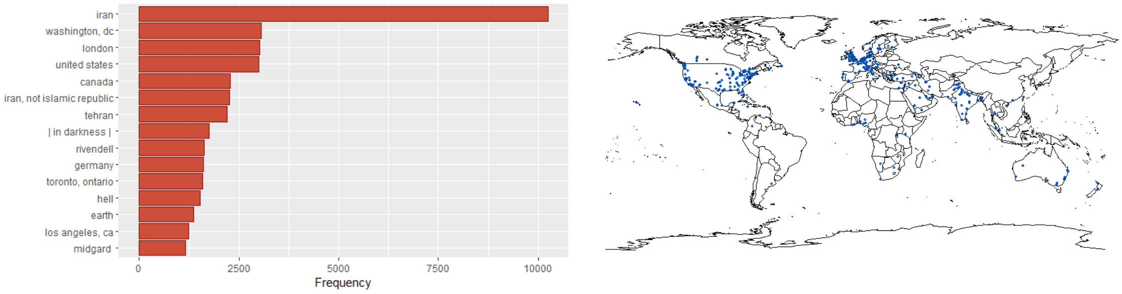


(a) January 2006 - January 2023

(b) September 2022 - October 2022

Figure 2: Frequency of created accounts per day

Some insight can also be drawn by looking at the user profiles. Figure 2 shows a histogram of the creation date of the accounts which have appeared in the dataset. It is possible to notice a sudden and extreme increase in the creation of Twitter profiles in the days following Mahsa Amini’s death, reaching the maximum exactly one week later, on 23 September 2022, with over a thousand profiles created during that day. As a matter of fact, 10766 users out of the 53335 unique users in the dataset have joined Twitter within a month of Amini’s death, i.e. the 20.2% of the total. Based on this data, it can be inferred that the users in question did not simply switch their conversations to the current socio-political hot topic, but rather actively joined the platform in order to participate in discussions surrounding the protests.



(a) Most common self-reported user locations

(b) Map of geo-located tweets

Figure 3

Figure 3a shows a histogram reporting the fifteen most common self-reported user locations along with a map containing the actual coordinates of tweets having geo-location tags, allowing us to get some information regarding the user-base demographics. An interesting fact is that, even though Iran is by far the most common self-reported location, the majority of self-selected locations in the dataset are situated in economically developed regions, specifically in Europe and North America. This is somewhat confirmed by the map of known geo-located tweets (Figure 3b), which shows a prevalence of tweets in those regions. This could be interpreted as a form of online activism: users from abroad change their location to Iran to show solidarity with the protests. However, only 1910 tweets in the dataset, or 0.3% of the total, carry geo-location metadata, meaning that these results must be considered with caution.

A couple more observations can be made about Figure 3a. First of all, there do appear some fictional or peculiar places, like “Rivendell”, “Midgard”, “hell” or “in darkness”. Moreover, the sixth most popular location is “Iran, not islamic republic”. This gives already an idea of the political stance associated with the userbase.

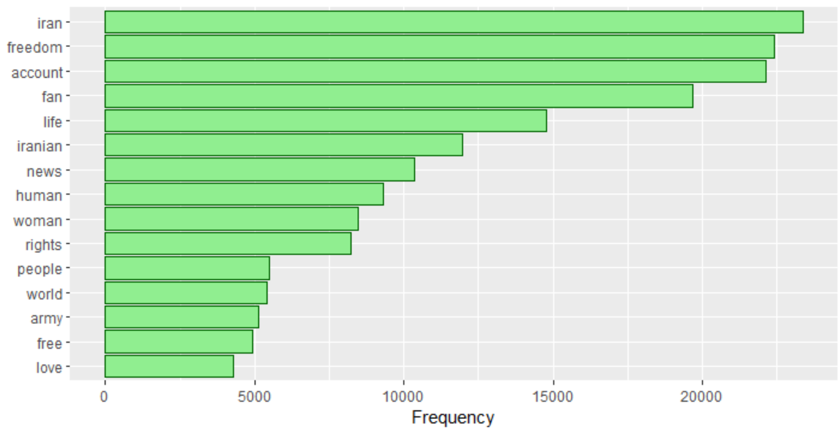


Figure 4: Frequency of the 15 most common words in in users’ bio

This is also confirmed by the most common words in users’ bios (Figure 4), where all three of the words associated with the signature slogan of the protests, "Woman, life, freedom" are present. One final consideration involves the presence of the word *news*, hinting at the fact that many of the scraped tweets could come from news agencies profiles.

This information allows us to sketch a profile of the users in the dataset: they mostly live in rich western countries, possess a strong internet connection and reside in urban areas, as evidenced by the high frequency of major cities in the locations.

3.2.2 Text Mining

To get a first understanding about what are the main topics of conversation around the protests, a histogram of the most common words is produced (Figure 5). It is already possible to notice a series of words, such as *sentenced, arrested, executed*, that almost certainly refer to the repercussions that some of the protesters in Iran faced, due to the harsh government crackdowns of the demonstrations. Moreover, there appear words like *please* and *voice*, which most likely point out an urgency to spread around the platform a message of help and solidarity.

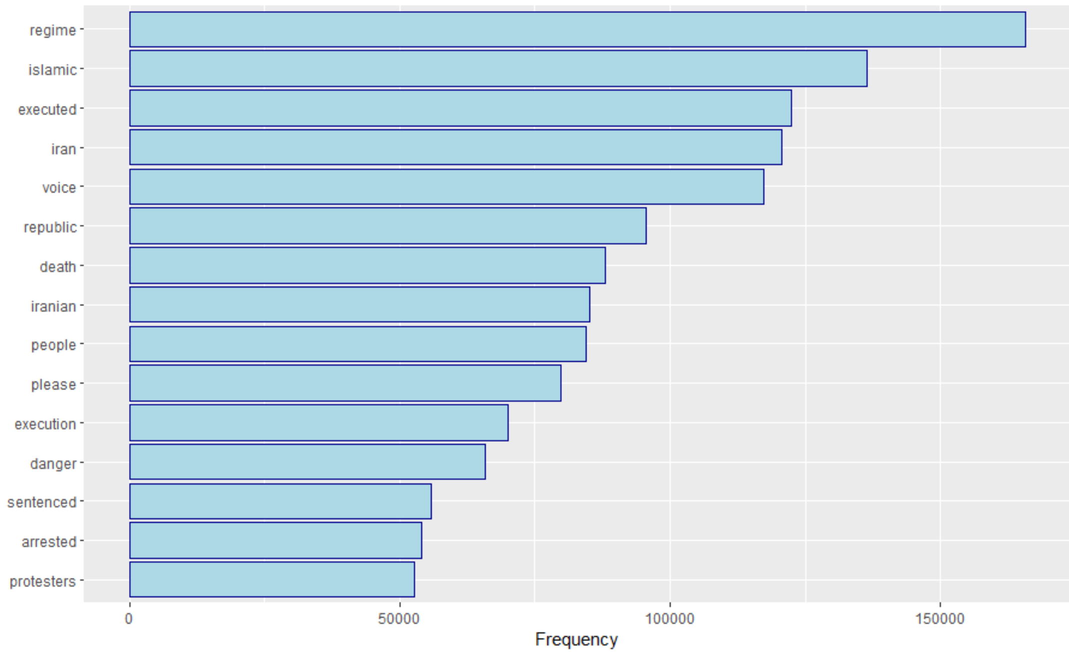


Figure 5: Frequency of the 15 most common words in the cleaned tweets

Then, a word network of the most common bigrams in the stemmed tweets is represented in Figure 6, allowing to identify patterns and trends in the tweets. In this case, only the bigrams appearing at least 10'000 times in the text are presented.

One peculiar bigram is the one that links *god* and *war*. As a matter of fact, it presumably derives from the English translation of *Moharebeh* (*i.e.* "waging war against God"), a criminal charge levied against people who commit acts against the Iranian government, including several demonstrators during the Mahsa Amini protests.

Another very frequent bigrams is *year-old*. One possible explanation for this focus on age may be that many of the convicted protesters are indeed in their early 20s [17].

The center of the conversation, however, is clearly on the protesters themselves, which are "sentenced to the death penalty", whose life is in "grave danger" and who suffer the "tortures" and "mass executions" carried out by the Islamic Republic.

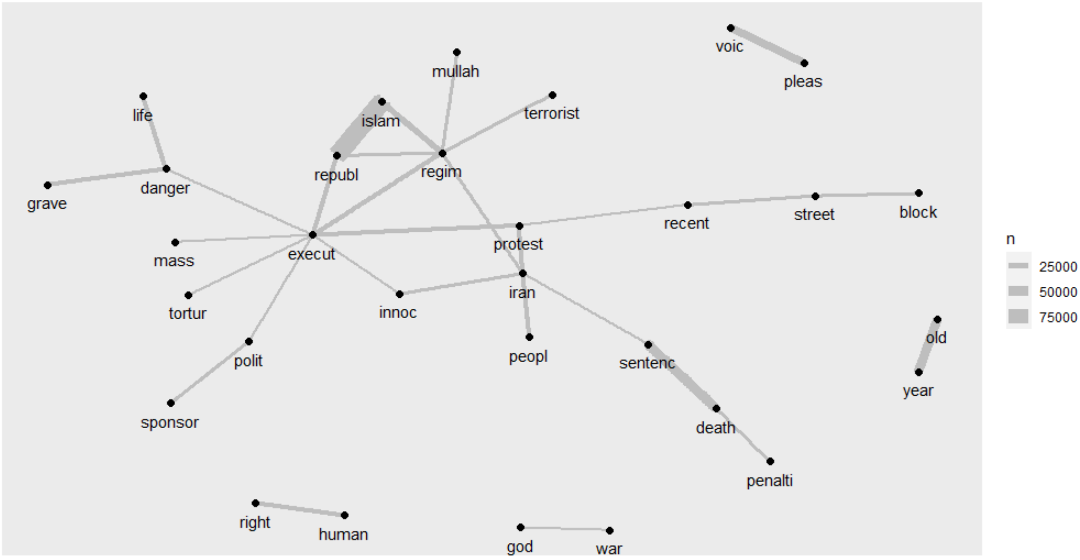


Figure 6: Bigram word network on the stemmed tweets

3.2.3 Topic Modelling

A powerful tool to identify topics in a corpus of texts is the Structural Topic Model (STM) [18]. It aims to understand how different topics are related to one another within a corpus, and how those relationships change over time, uncovering patterns and trends in the data.

Due to its high computational cost, the algorithm is not run on the whole dataset, but instead over a subset of 50'000 randomly sampled tweets having at least 6 word tokens. The algorithm maximizes semantic coherence with a number of topics equal to 12, so this number was used for the analysis.

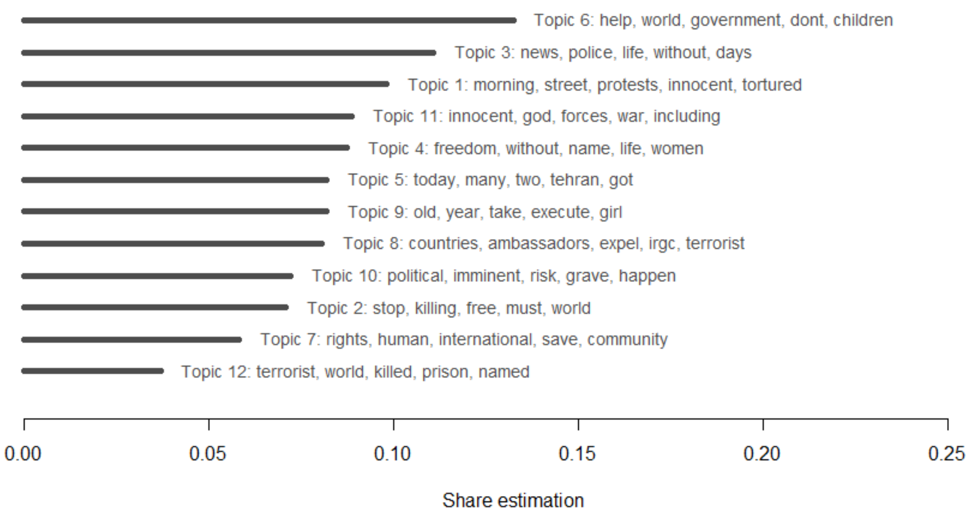


Figure 7: Shares of each topic according to STM

Figure 7 identifies the topics' respective total share, along with five representative words for each. It is possible to notice that the most prevalent is Topic 6, which is associated to words such as *help, world, government*. This, along with Topic 2 and 7, seems to center around a call for help to the international community and human rights organizations. Topic 8 seems to possibly hint at a specific request, namely for countries to expel their Iranian ambassadors. Topic 1 and 3 are more focused on giving news and information describing the protests, whereas Topics 5 and 9 appear to be mostly capturing common words.

In Figure 8, the relationship between the week of year in which the tweet was created and the six most common topics is plotted. For example, tweets created until Sunday 11 December at 23:59 belong to week 1, whereas tweets created between Monday 12 December at 00:00 and Sunday 18 December at 23:59 belong to week 2, and so on. In this case, the week variable is allowed to have a non-linear relationship in the topic estimation stage.

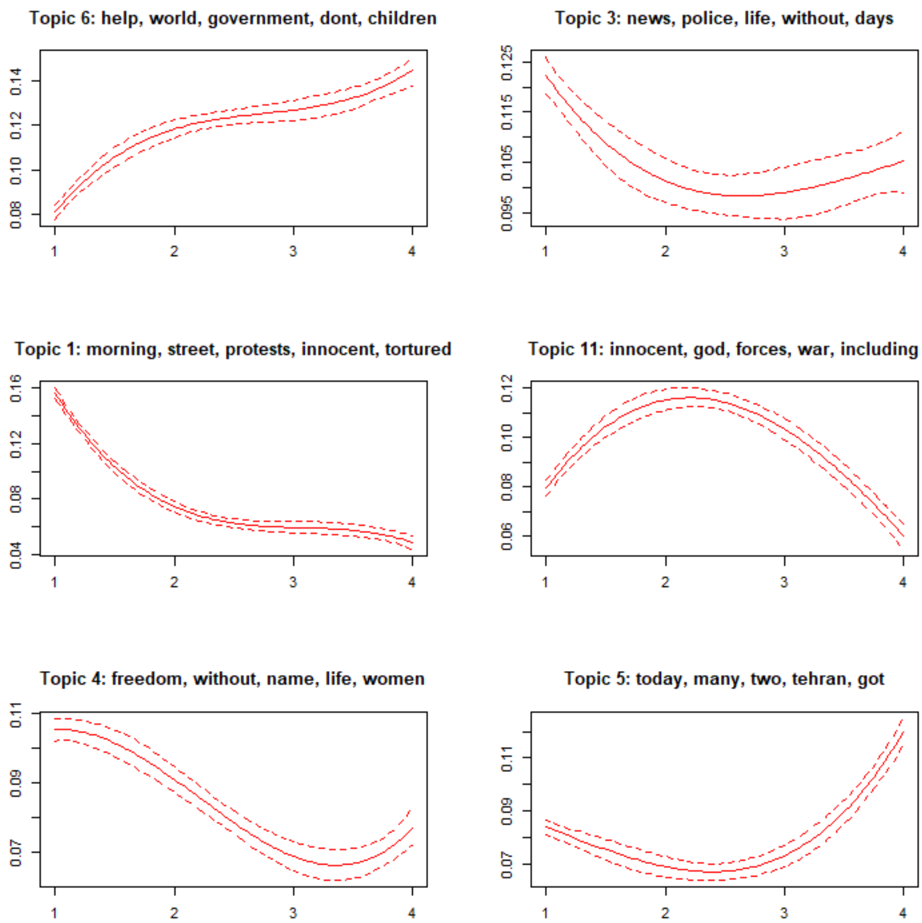


Figure 8: Shares of most frequent STM topics over time

Looking at the left column, it is possible to notice that Topic 6 is positively associated with the week variable, whereas Topic 1 and 4 have a decreasing trend. Topic 1 seems focused about describing what happened to the protesters (*innocent, protests, tortured*), whereas Topic 4 contains the three words of the protests' slogan (*women, life, freedom*). This could hint at the fact that, in the first period, people were mostly tweeting about the first executions of the protesters, using the slogan very often in their tweets. As time went by, the focus of the conversation shifted more onto a plea for help, as exemplified by Topic 6.

Considering the right column, Topic 3 and 5 have a "U-shaped" relationship, meaning that they were more prevalent during week 1 and 4, and less prevalent during week 2 and 3. The inverse is true for Topic 11. In this case, however, interpreting these relationships is not straightforward, given the presence of many common words captured by these topics.

3.2.4 Sentiment Analysis

The final step of the procedure involves performing a Sentiment Analysis on the tweets.

Figure 9 compares the most common words for positive and negative sentiment, according to the *Bing* sentiment lexicon [19]. It can be noted how the words *death* and *danger* are extremely common, appearing respectively roughly three times and two times more than the third most frequent negative word, *crime*. Moreover, it is possible to notice how the frequencies of the most common positive words are much lower than their negative counterparts.

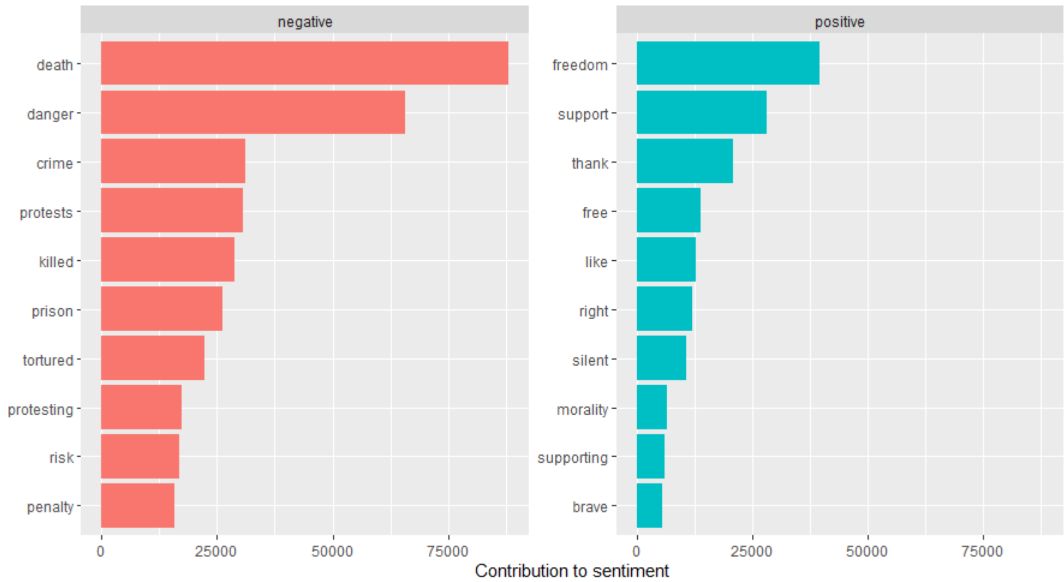


Figure 9: Most common words for positive and negative sentiment according to the Bing lexicon

The tweets' sentiment is then analyzed over the whole time period in question, using a lexicon-based approach with three different dictionaries, namely *Syuzhet, AFINN* and *Bing*. When performing this kind of sentiment analysis, the text is first tokenized, *i.e.* broken down into individual words. Then, each word is looked up in the lexicon to see if it is associated with a particular sentiment. The three lexicons have different scales: *Syuzhet* and *Bing* assign to each word a value in the [-1,1] range, whereas AFINN uses a [-5,5] range. Results are shown respectively in Figures 10, 11 and 12.

It is possible to notice that the trend of the first two charts is very similar: there are two local minima around the midday of December 8 and 11, followed by a peak on December 15 and another trough around December 19. From then onwards, the sentiment seems to increase, with a slight decline between December 24 and 27.

The third graph is somewhat different, having no minimum on December 8 and having a trough on December 27. However, it confirms the two sentiment lowest points around December 11 and 19.

Overall, it is worth noting that, according to all three methods, the sentiment is always negative.

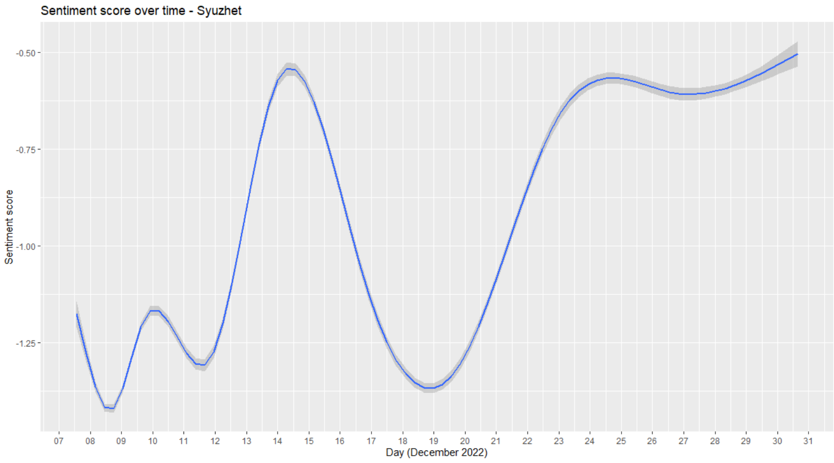


Figure 10: Sentiment over time according to Syuzhet

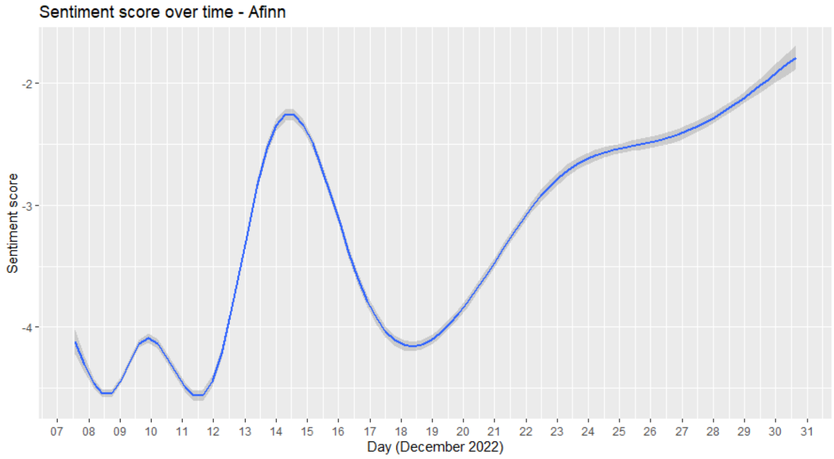


Figure 11: Sentiment over time according to AFINN

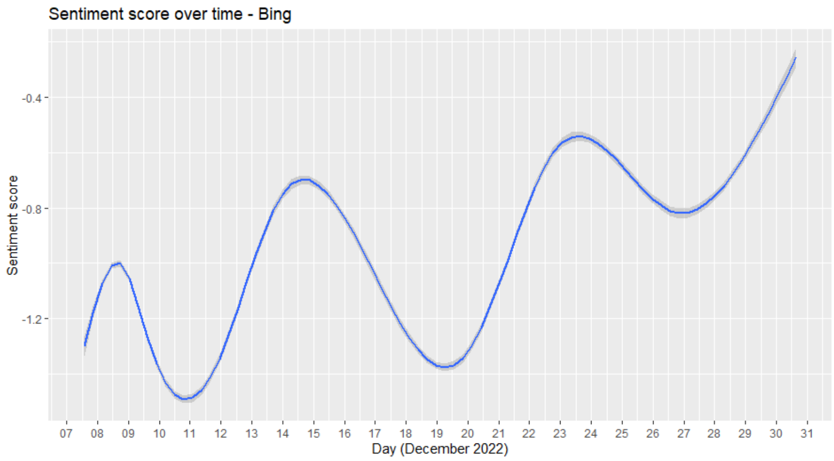


Figure 12: Sentiment over time according to Bing

Another useful tool is the NRC sentiment lexicon, which, instead of only focusing on whether words are positive or negative, categorizes text using 8 additional types: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust [20]. The frequency of each emotion found in the text is calculated, and the result is shown in Figure 13. The most frequent emotion is by far fear, appearing over 71'000 times, followed by anger and sadness, with a score slightly over 50'000.

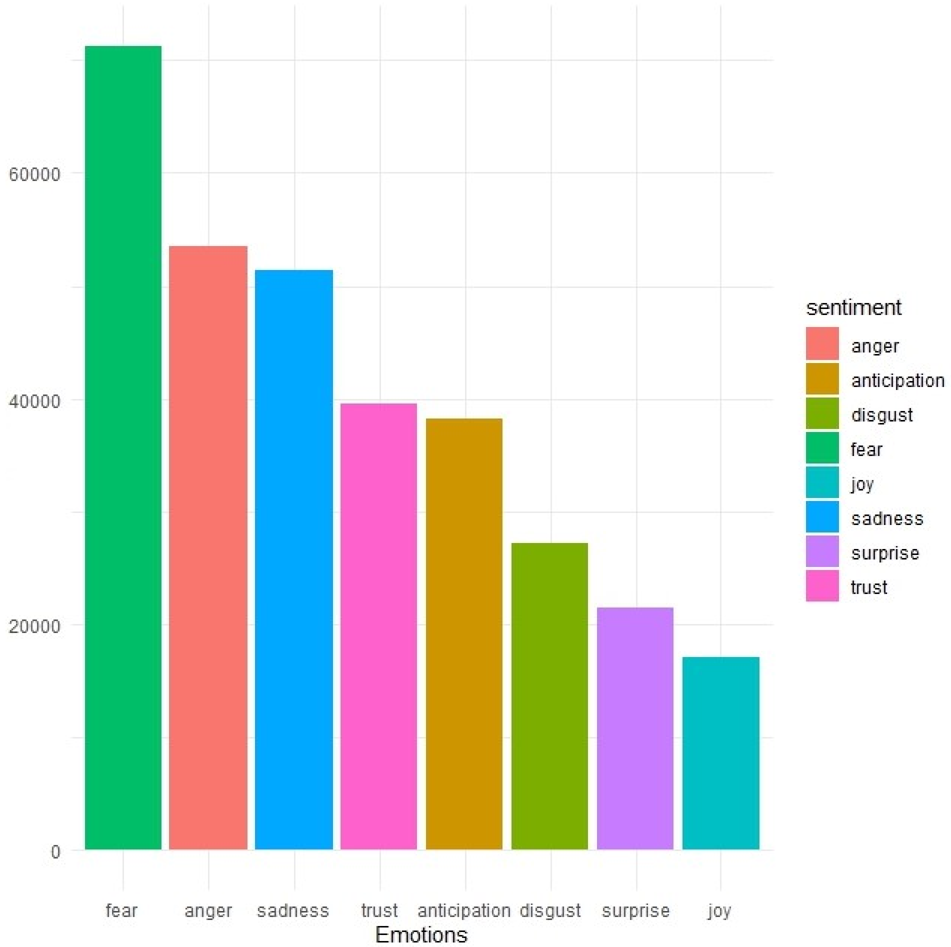


Figure 13: Histogram of emotion scores according to the NRC lexicon

It is also possible to look at the evolution of the prevalence of the NRC emotions over time. In particular, Figure 14 focuses on the percentage of tweets expressing the three most prevalent emotions in the dataset: fear, anger and sadness. It is clearly possible to see how fear and sadness seem to follow the same trend, albeit with a somewhat constant difference in the percentage of tweets of around 3-5%. Moreover, the spikes in these two emotions correspond to the local minima in Figure 10 and 11 and to the three periods with the highest volume of tweets, as identified in 3.2.1. This is unsurprising, given that those periods coincided with executions of protesters.

Anger also seems to initially follow fear's trend, but after 13 December, the two decouple. However, it can be observed that after 20 December, the percentage of tweets expressing fear decreases, whereas the prevalence of anger increases, arriving at a point in which they are almost equal on 30 December.

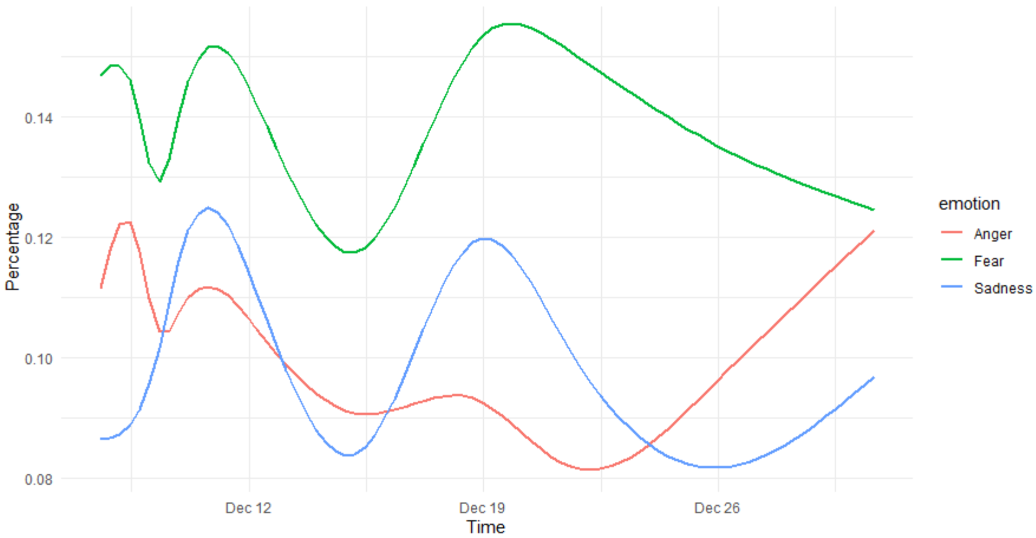


Figure 14: Anger, sadness and fear in tweets over time according to NRC

4 Conclusion

Overall, the employed techniques permit to obtain a series of key insights into the Twitter conversation about the current protests in Iran. First of all, the evidence gathered so far suggests a strong component of urban, western non-iranians generating awareness about the situation. Moreover, the fact that one fifth of the users in the dataset have joined Twitter right as the protests gained momentum could suggests that they joined the platform specifically to be involved in them.

The text mining step revealed that the most frequent words used in tweets about the Iran protests were related to human rights, executions and calls for help to the international community. The topic modelling analysis corroborated these results and allowed to capture the temporal trend of the prevailing topics of discussion during the studied time frame.

The results of the sentiment analysis showed that negative tweets are much more prevalent than the positive ones, and that, over time, the sentiment reflected key moments in the demonstrations, namely the first executions of the protesters. Moreover, fear appeared to be the most frequent emotion in the tweets, followed by anger and sadness.

This analysis focused on the textual features and language employed by Twitter users who tweeted about the Mahsa Amini protests between 7 and 30 December 2022. However, there are many other aspects that could have been studied to gain a more comprehensive understanding of the role of Twitter in political activism and protest movements. For instance, network analysis can provide useful information to identify key influencers and leaders within the movement, as well as to track the spread of information and ideas by analyzing the relationships between users and specific hashtags [8]. Furthermore, investigating the images, videos and links shared on the platform could also provide another level of depth to the analysis [21].

Additionally, it is worth mentioning that the study considered only tweets produced during a short time frame. It could be suggested to further deepen it by considering tweets from the start of the protests in September 2022, allowing to track the emergence of the movement as the first protests gained momentum.

Finally, Twitter data could be combined with other data sources, such as other social media platforms, traditional news articles and surveys, in order to gain a more comprehensive understanding of the protests' dynamics, messages and participants.

References

[1] Rana Rahimpour. Fury in iran as young woman dies following morality police arrest. `https://www.bbc.com/news/world-middle-east-62930425`, BBC News, September 16, 2022. Accessed on January 11, 2023.

[2] Iran Human Rights (IHRNGO). At least 100 protesters facing execution, death penalty charges or sentences; at least 476 protesters killed. `https://iranhr.net/en/articles/5669/`, December 27, 2022. Accessed on January 11, 2023.

[3] Vivian Yee and Farnaz Fassihi. 'out-of-reach dreams' in a sickly economy provoke the rage in iran. `https://www.nytimes.com/2022/10/02/world/middleeast/iran-protests-economy.html,TheNewYorkTimes`, October 2, 2022. Accessed on January 13, 2023.

[4] Shabnam von Hein. Iranians use social media to keep protest movement alive. `https://www.dw.com/en/iranians-use-social-media-to-keep-protest-movement-alive/a-63767075,DeutscheWelle`, November 15, 2022. Accessed on January 13, 2023.

[5] Evgeny Morozov. Iran: Downside to the" twitter revolution". *Dissent*, 56(4):10–14, 2009.

[6] Alok Choudhary, William Hendrix, Kathy Lee, Diana Palsetia, and Wei-Keng Liao. Social media evolution of the egyptian revolution. *Communications of the ACM*, 55(5):74–80, 2012.

[7] Li Tan, Suma Ponnam, Patrick Gillham, Bob Edwards, and Erik Johnson. Analyzing the impact of social media on social movements: a computational study on twitter and the occupy wall street movement. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 1259–1266, 2013.

[8] Mark Tremayne. Anatomy of protest in the digital era: A network analysis of twitter and occupy wall street. *Social movement studies*, 13(1):110–126, 2014.

[9] Devin Gaffney. iranelection: Quantifying online activism. In *In Proceedings of the Web Science Conference (WebSci10*. Citeseer, 2010.

[10] Lev Grossman. Iran protests: Twitter, the medium of the movement. *Time Magazine*, 17, 2009.

[11] Onook Oh, Chanyoung Eom, and H Raghav Rao. Research note—role of social media in social change: An analysis of collective sense making during the 2011 egypt revolution. *Information Systems Research*, 26(1):210–223, 2015.

[12] Gema Bello-Orgaz, Jason J Jung, and David Camacho. Social big data: Recent achievements and new challenges. *Information Fusion*, 28:45–59, 2016.

[13] Justin Grimmer, Margaret E Roberts, and Brandon M Stewart. *Text as data: A new framework for machine learning and the social sciences*, chapter 3. Princeton University Press, 2022.

[14] David Gritten. Mohsen shekari: Iran carries out first execution over protests. `https://www.bbc.com/news/world-middle-east-63900099`, BBC News, December, 8 2022. Accessed on January 14, 2023.

[15] Parisa Hafezi. Iran carries out second execution linked to wave of popular protests. `https://www.reuters.com/world/middle-east/iran-carries-out-second-execution-linked-anti-government-protests-2022-12-12/`, Reuters, December 20, 2022. [Accessed on January 21, 2023].

[16] Maryam Afshang. Iran executes 2 men arrested in protests. https://www.nytimes.com/2023/01/07/world/middleeast/iran-executes-protesters.html, The New York Times, December 7, 2023. [Accessed on January 21, 2023].

[17] Farnaz Fassihi. Stymied by protests, iran unleashes its wrath on its youth. https://www.nytimes.com/2022/11/14/world/middleeast/iran-protests-children.html, The New York Times, November 14, 2022. [Accessed on January 20, 2023].

[18] Margaret E Roberts, Brandon M Stewart, Dustin Tingley, Edoardo M Airoldi, et al. The structural topic model and applied social science. In *Advances in neural information processing systems workshop on topic models: computation, application, and evaluation*, volume 4, pages 1–20. Harrahs and Harveys, Lake Tahoe, 2013.

[19] Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177, 2004.

[20] Saif M Mohammad and Peter D Turney. Nrc emotion lexicon. *National Research Council, Canada*, 2:234, 2013.

[21] Hyunjin Seo. Visual propaganda in the age of social media: An empirical analysis of twitter images during the 2012 israeli–hamas conflict. *Visual Communication Quarterly*, 21(3):150–161, 2014.