**Sentiment and Frame Analysis: A Ukraine-Russia Dispute Study**

**Rushabh Khara**

**Introduction**

Vladimir Putin announced anti-submarine drills with ally Belarus near the Ukrainian border on February 14, 2022. The US made predictions and issued warnings about potential invasions. Defending the actions as "military drills," the Russian president made a spectacle of the accusations. Through the course of events, on 24th February, Vladimir Putin sanctioned a full-fledged war against Ukraine to annex former parts of the Soviet Union. Throughout the globe, governments condemned the decision, but there has been little study about the change in public sentiment around the issue as time passes. Sentiment analysis allows us to study public opinions, feelings, and attitudes toward any event, person, or topic (Venugopalan & Gupta, 2015). This study proposes a combination of frame and sentiment analysis techniques to uncover the significant themes and corresponding public attitudes and provide interpretation in small sentences. This method can help identify notable events around which political parties/governments can build agendas to gain public support. At a personal level, party leaders might be concerned that critical public sentiment and political backlash might lead to defeat in the upcoming election (Tomz et al., 2019). Further, MNCs can try to side with favourable entities to catch the public eye. There can be several other implications; the above are just some possible examples. The emergence of microblogging platforms has transformed them into a sentiment-rich data source. Our focus remains on social media platforms due to their more expansive reach than traditional media during the crisis (Garcia & Cunanan-Yabut, 2022). We have chosen Twitter for its popularity worldwide as a source of information exchange.

**Literature Review & Research Questions**

Go et al. (2019) are credited with one of the earliest attempts at extracting sentiment from Twitter. There have been various studies on sentiment analysis, but they use computationally heavier techniques like Naive Bayes and Machine Learning. Garcia & Cunanan-Yabut (2022) have formerly studied public opinion around the Ukraine-Russia war. We propose to build upon and provide a modified approach integrated with frame analysis to gain novel insights. A multidisciplinary social science research technique, frame analysis, examines how people perceive situations and activities. The first instances of Frame Analysis date back to 1974 under the work of Erving Goffman. Miller (1997) focused on conservationists and property owners while examining the issue frames surrounding wetlands in the US. We will use Miller's suggested method to include as many keywords in a frame to create informative statements.

The paper aims to answer the following research questions:

1. What is the current public sentiment around the Ukraine-Russia war?
2. What are the discourse's major themes/events/topics?
3. Is one country supported more than the other country?

**Data & Methods**

The data has been collected using VOSONDash with API keys for Twitter created using the help of Dr. Robert Ackland. A total of 2000 tweets have been retrieved using hashtags (*“#RussiaUkraine,” “#UkraineRussia,” “RussiaUkraineWar,” and “UkraineRussiaWar”*) that have been selected to avoid bias towards any particular country. For example, *“#StandWithUkraine”* would create an obvious bias and lead to incorrect inferences. We should note that only recent tweets have been selected, which might lead to different insights than Garcia & Cunanan-Yabut (2022). To deal with synonyms, a dictionary is used, which reduces the number of features and significantly affects the relative importance of keywords.

All the analyses have been performed using *Gephi* and *R Programming* in *RStudio.* The extracted tweets have been pre-processed by removing symbols, punctuations, numbers, and mentions. Tweets contain emojis/emoticons that are not removed by the *tokens()* function, which we suspect is due to the encoding error. Custom words have been used with *stopwords()* to clean data and remove general keywords. Further, we transformed the data into a data frame and document feature matrix after tokenisation based on the analysis to be conducted. Sentiment analysis has been supplemented using the Lexicoder Sentiment Dictionary (Young & Soroka, 2015), which classifies the token’s polarity as “neutral,” “positive,” or “negative.”

A bipartite network between co-occurring keywords has been created to analyse frames, where nodes represent themes/concepts, and ties between them indicate co-occurrence in a tweet. Visualisations are created with the *worldcloud2* library, *wordcloud* library, and *Gephi,* where keywords with a frequency higher than five are only projected. The size of words in all the visualisations indicates their respective frequency, except the frame analysis, where size corresponds to a keyword's degree of centrality.

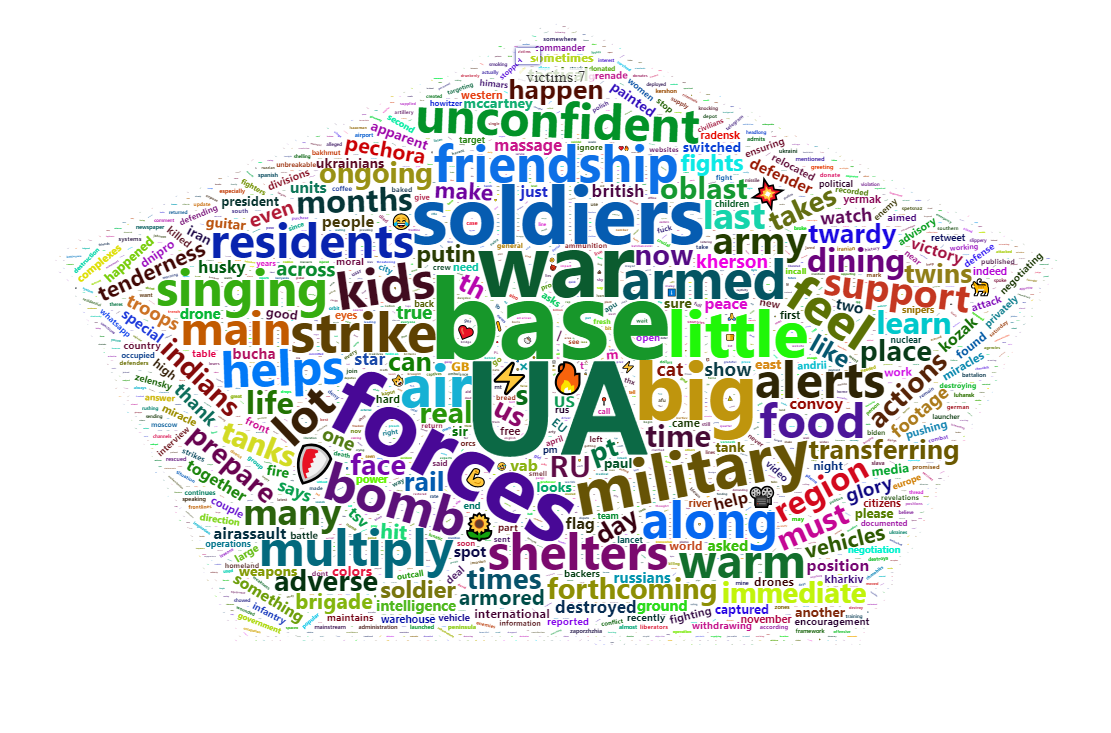
**Results & Discussion**

Fig. 1. Most frequent words

Glancing at Fig.1 and Table 1, we can observe the most frequent keywords used during the Ukraine-Russia discourse, excluding “Ukraine,” “Russia,” “Ukrainian,” and “Russian.” The words have been omitted as they are of no significance in interpreting the themes. Top words do not show any sign of sentiment, opposite to what Garcia & Cunanan-Yabut (2022) discovered, but are more indicative of prominent themes in the war in the current scenario. The difference might be because the war started several months ago, and now the feelings might have subsided, which can be interpreted as human nature, where the intensity of emotions decreases as time passes.

| Word | Frequency | Tweets |
| --- | --- | --- |
| UA | 319 | "A russian mobik hiding in the woods was captured by the UA" |
| base | 201 | "The main residents of the military base.Because the base has a lot of food and soldiers to warm up." |
| forces | 165 | "A senior commander of the notorious mercenary group has been killed in an ambush by special forces. Pavlo Pataretsky, the Khort commander spoke of the incident. reports➡️ " |
| war | 161 | "Iran confirms drones to Russia but ‘months’ before Ukraine war." |
| big | 143 | "Let's see which Countries support Nazism.Ohho a big number . Wait What, all are actually country." |
| soldiers | 143 | "Afghan special forces soldiers who fought alongside American troops and then fled to Iran after the chaotic US withdrawal last year are now being recruited by the Russian military to fight in Ukraine." |
| little | 115 | “Ministry of Defence said K conscripts in will have little impact on bcoz is struggling to train them" |
| military | 115 | "Ukrainian troops use residential buildings for military purposes.” |
| air | 104 | "During multiply air strike alerts not only Ukrainian kids feel unconfident in the bomb shelters. So singing along helps." |
| strike | 100 | "During the day, the Air Force of the Armed Forces of Ukraine launched three strikes on the enemy. The missile and artillery units hit nine enemy personnel, weapons and military equipment clusters and an ammunition depot." |

Table. 1. Frequency of top keywords with relevant tweets.

The top keywords are all related to recent events during the war, indicating the real-time information being conveyed via Twitter. The initial frequency and tweet analysis performed above explain the prominent themes. The high frequency of words like “soldiers,” “strike,” and “war” makes sense due to the destructive state of the situation. Bigrams have not been utilised due to their nature of projecting general keyword pairs like “for you,” “of the,” and “not only,” which populate the whole word cloud, and we, as researchers, miss out on important thematic insights.

In contrast to keywords, hashtags are a concise way to show partisanship and sentiment towards a particular entity/topic, which makes them the perfect choice for sentiment analysis. Hashtags can be used to comprehend the sentiment towards an entity and uncover whether support for two entities differs. Albina et al. (2019) suggest hashtag classification can be up to 95% accurate if combined with SVM and Naive Bayes algorithms and up to 80% using the standard method. For this paper, we will stick to standard classification and verify the credibility of the claims.

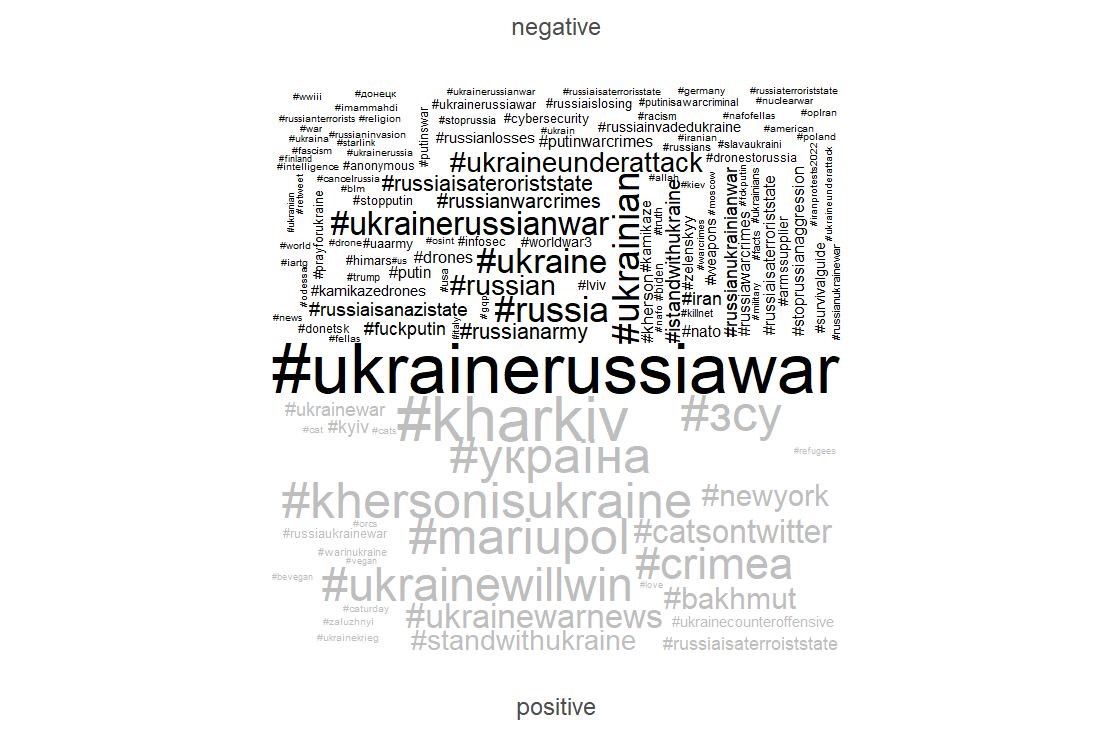


Fig. 2. Hashtag Sentiment Word Cloud

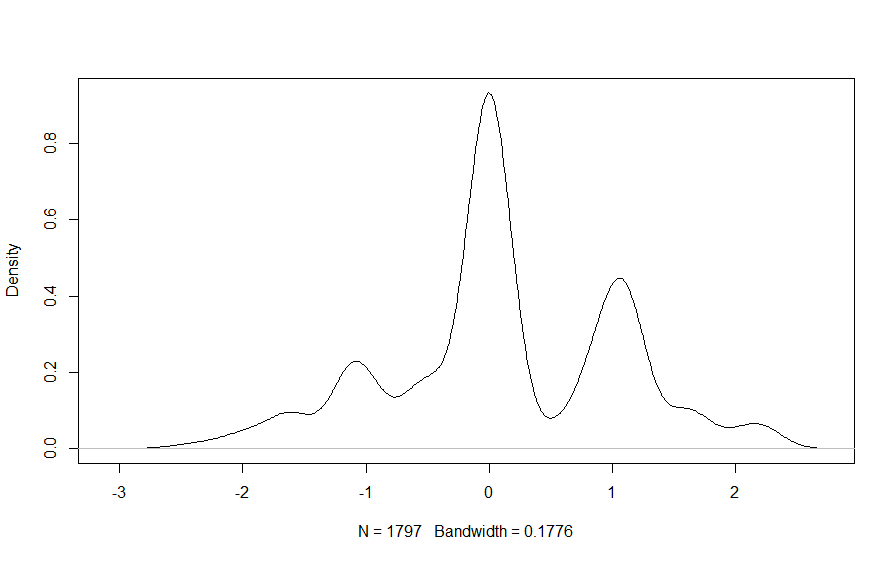
Strong negative sentiments against Russia can be seen in Fig.2 (*“#fuckputin,” “#putinwarcrimes,” “#russiaterroriststate,” “#russianwarcrimes”*), which resonate with the sentiment put forth by the majority of countries. The hashtags show hate feelings and condemnation towards actions taken by Vladimir Putin. However, increasingly positive/sympathetic hashtags like *“#ukrainewillwin”* and *“#standwithukraine”* show support towards Ukraine. The word cloud does not show any negative sentiment word directed toward Ukraine. In Fig.3, we observe a higher density of positive keywords than negative ones, which is inconsistent with the findings of Garcia & Cunanan-Yabut (2022), where a higher negative keyword density was observed. It is suspected that over time negative sentiment reduces and is replaced by positive sentiment. Anger is an impulsive emotion and immediate reaction to something wrong; however, over time, it gets substituted by empathy/sympathy and supportive emotions. Table 2 shows tweets containing some of the hashtags from Fig.2. Tweets help us to look at keywords in context. We could have used the *kwic()* function for a similar purpose. However, it sometimes provides a vague understanding of the sentiment by only providing a few neighbouring keywords, making it difficult to verify the correctness of the sentiment classification.

Fig. 3. Sentiment Density Plot (Logit Scale Sentiment Score)

| *#russiaisateroriststate* | *"One man's perilous mercy mission to ease plight of civilians on frontline in Kherson - Sky News\n #UkraineRussiaWar #Ukraine #RussianUkrainianWar #RussianWarCrimes #StandWithUkraine* ***#russiaisateroriststate*** *#UkraineWar #UkraineWillWin #UkraineUnderAttack"* |
| --- | --- |
| *#ukrainecounteroffensive* | *"Between fights there is time for tenderness.\n\n#Ukraine️ #Ukraina #RussiaisATerroistState #RussiaUkraineWar #CatsOfTwitter #CatsOnTwitter #Bakhmut #Kharkiv\n #NewYork #Zaporizhzia #Kyiv* ***#ukrainecounteroffensive*** *https://t.co/LIfm2izhrQ"* |
| *#standwithukraine* | *"Ukrainian drone dropped a bomb on russian occupant. New video. Somewhere in Ukraine 🇺🇦 \n#Ukraine #UkraineRussiaWar #UkraineRussianWar #RussiaIsLosing* ***#StandWithUkraine️*** *https://t.co/nY4KMkeUix"* |
| *#russianwarcrimes* | *"Dead russian soldier\n #UkraineUnderAttack #Ukraine #UkraineWillWin* ***#RussianWarCrimes*** *#StandWithUkraine #RussianUkrainianWar #UkraineRussiaWar #russiaisateroriststate #UkraineWar https://t.co/XkViy3xYu6"* |

Table 2. Tweets related to hashtags.

In Fig.2, most hashtags include the word *“Ukraine”* and only depict negative sentiment against the Russian side. It does not mean that there are no Russian supporters but that the support for Ukraine overshadows it. Understanding the major events around the war is crucial, but reading every tweet is impractical; hence we analyse frames that serve as a tool to analyse the influence of events/entities on society in short sentences. Frame analysis is conducted on *Gephi* instead of *Rstudio* due to its superior visualisation properties. Gephi helps with swift exploration and the ability to display large networks in real-time using a 3D renderer.

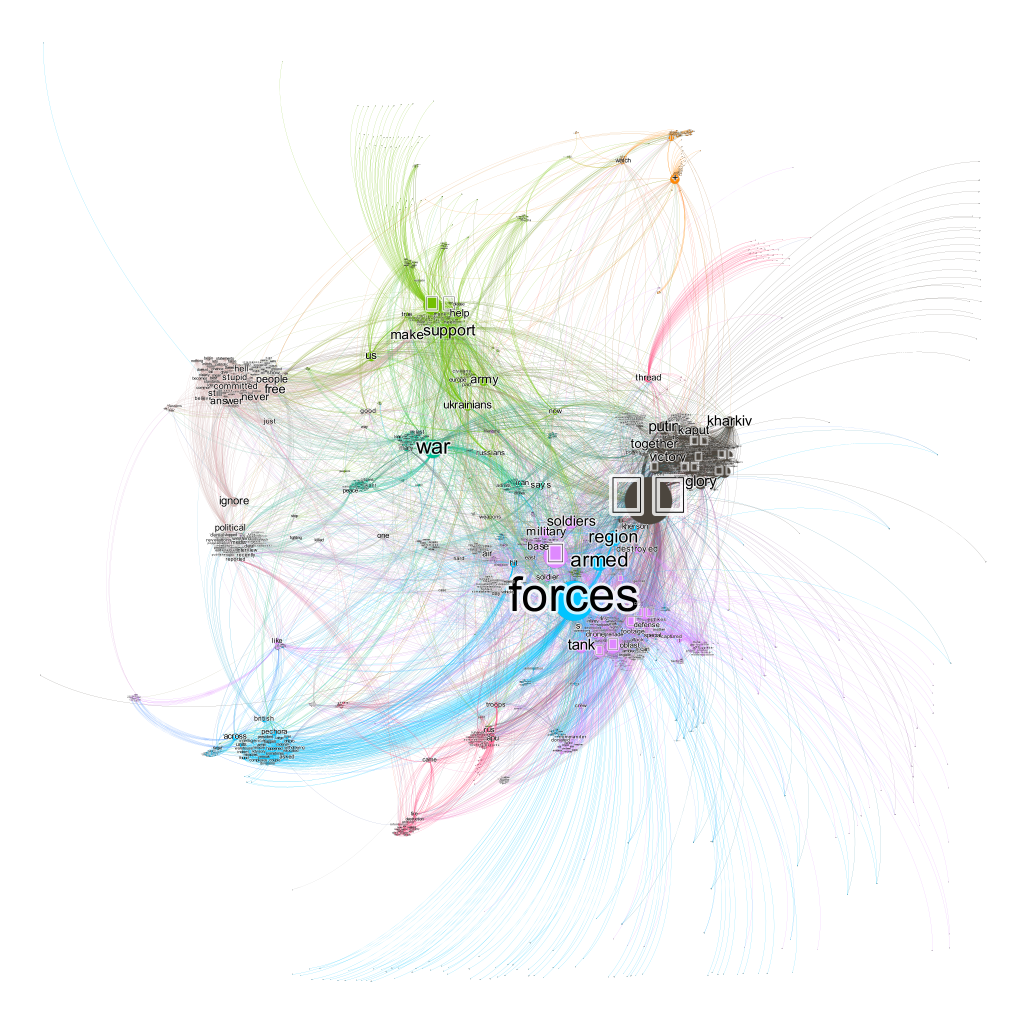


Fig. 4. Frame network created on Gephi.

Empty boxes appearing in Fig.4 represent emojis/emoticons that *Gephi* cannot interpret due to encoding issues. We could not remove the emojis during the data preprocessing, which indicates methodological issues in our approach that are open to improvement.

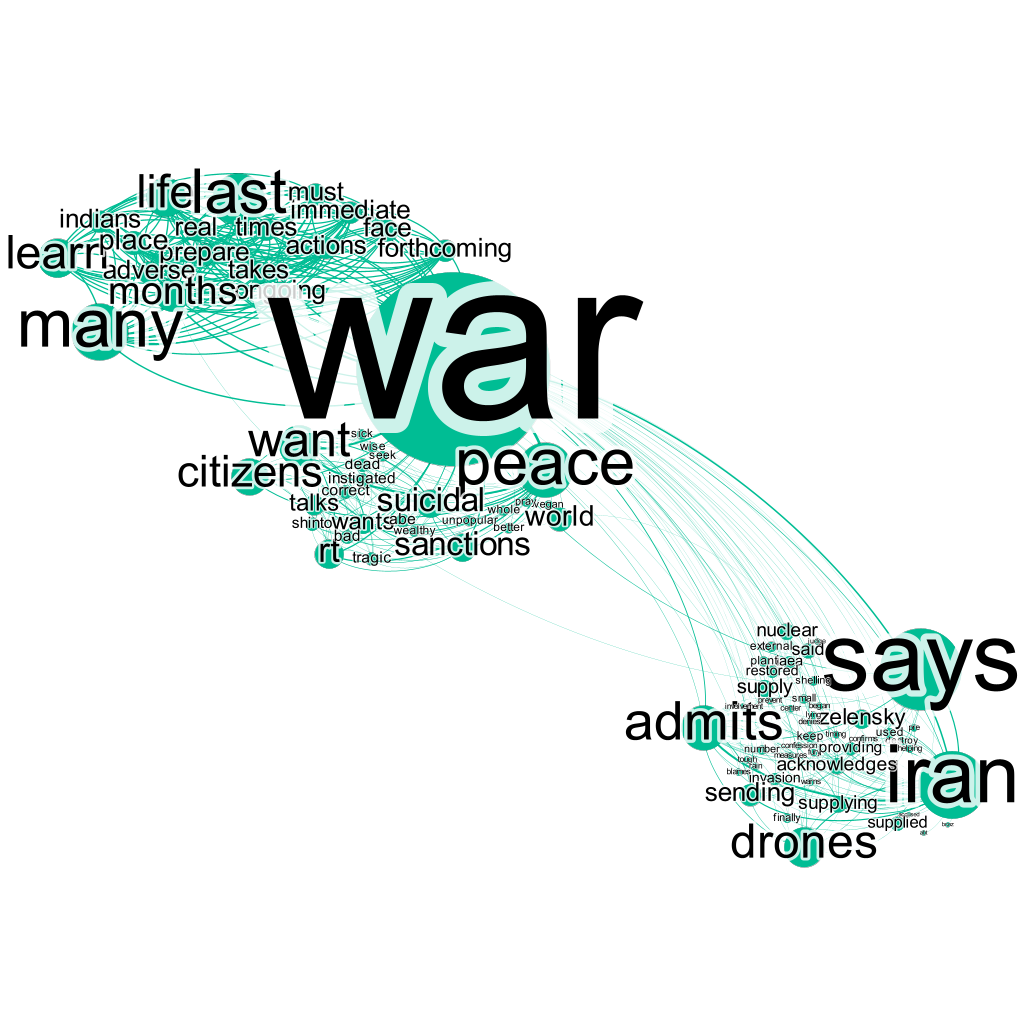


Fig. 5. Frame 1

| *Frame 1:* ***Iran admits*** *to providing* ***drones*** *to Russia,* ***says*** *it was* ***many months*** *before the* ***war*** *started, and supports* ***peace****.* |
| --- |

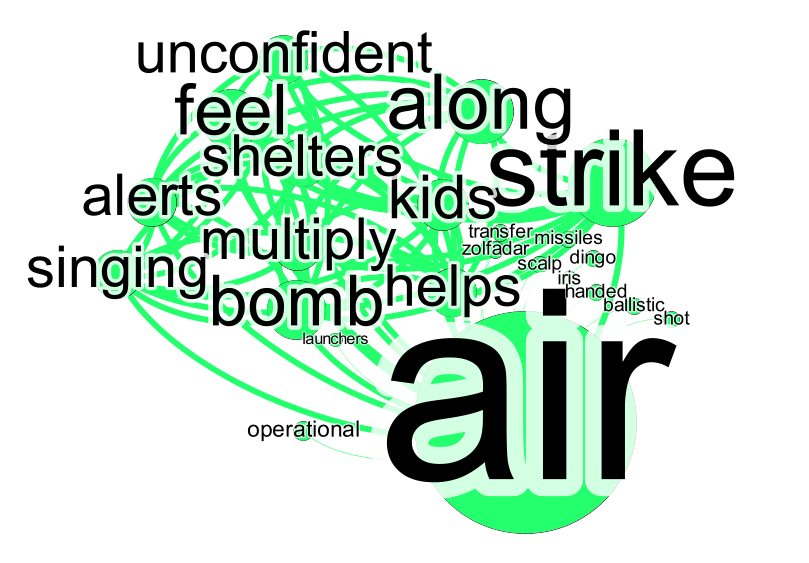


Fig. 6. Frame 2

| *Frame 2: During* ***multiply air strike alerts,*** *not only Ukrainian* ***kids feel unconfident*** *in the* ***bomb shelters****. So* ***singing along helps.*** |
| --- |

Frames in Fig.5 and Fig.6 reiterate the findings from Table 1 and manual tweet analysis, proving the method's high accuracy. We have followed the method Miller (1997) suggested and included as many keywords in a frame as possible to create an informative sentence. Entman (1991) specifies that framing involves “selecting and highlighting some features of while omitting others,” which is why we omit non-useful tokens like “launchers,” “iris,” “transfer,” and many more.

In frame analysis, we typically start with a specific thematic framework defining a specific belief or experience, then attempt to identify several related themes and the subtopics that serve as their tokens (Biria, 2016). However, frame analysis does not allow the complete interpretation of an event unless it is supplemented by appropriate domain knowledge. Researchers should carefully select *stopwords,* as removing words that might not seem helpful at first can alter the meaning of a frame or remove certain information.

We can back this claim by looking at Fig.6, where it is not apparent that kids from which country are singing, as we removed the keywords “Ukraine” and “Russia.” It was identifiable due to the domain knowledge at our exposure and thorough research on the topic.

Using Twitter data comes with its drawbacks. The microblogging platform has limited a tweet to 280 characters which do not provide enough conclusive information to analyse themes, frames, and sentiments. Another difficulty in overcoming is the prevalence of misspellings and slang words in tweets, which is much higher than in other language resources (Venugopalan & Gupta, 2015).

**Conclusion**

A hybrid model for classifying the sentiment and frames from tweets that incorporates unigrams and domain-specific lexicons has been modified and used to gain insights. In the study, we investigated the public sentiment around the Ukraine-Russia war, uncovering some crucial events/issues that were prominent during the discourse.

Additionally, it was discovered that negative sentiment was projected towards Russia compared to a sympathetic sentiment towards Ukraine alongside substantially high support. We discover that higher positive sentiment is depicted during the discourse, which we suspected is due to the sympathy and supportive keywords for Ukraine. Although the sentiment towards Russia is highly negative, the overall sentiment of the discourse is found to be more skewed toward the positive side. Frames analysed showcase some significant events that caught the public eye during the war while reiterating our findings in Table 1.

The work can be improved by utilising more advanced algorithms and techniques, modifying the preprocessing to eliminate the emoji/emoticons, or utilising an encoding that can interpret various symbols to facilitate better visualisations.

**References**

[1] Alfina, I. *et al.* (2017) “Utilizing hashtags for sentiment analysis of tweets in the political domain,” *Proceedings of the 9th International Conference on Machine Learning and Computing* [Preprint]. Available at: <https://doi.org/10.1145/3055635.3056631>.

[2] Entman, R. M. (1991). Framing U S coverage of international news: Contrasts in narratives of the KAL and Iran air incidents. Journal of Communication, 43(4), 51-58.]

[3] Garcia, M.B. and Cunanan-Yabut, A. (2022) “Public sentiment and emotion analyses of Twitter data on the 2022 Russian invasion of Ukraine,” *2022 9th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)* [Preprint]. Available at: <https://doi.org/10.1109/icitacee55701.2022.9924136>.

[4] Go, Alec & Bhayani, Richa & Huang, Lei. (2009). Twitter sentiment classification using distant supervision. Processing. 150.

[5] Miller, M.M. (1997) “Frame mapping and analysis of news coverage of contentious issues,” *Social Science Computer Review*, 15(4), pp. 367–378. Available at: <https://doi.org/10.1177/089443939701500403>.

[6] Smirnova, A., Laranetto, H. and Kolenda, N. (2017) “Ideology through sentiment analysis: A changing perspective on Russia and Islam in NYT,” *Discourse & Communication*, 11(3), pp. 296–313. Available at: <https://doi.org/10.1177/1750481317699347>.

[7] Soroka, S., Young, L. and Balmas, M. (2015) “Bad news or mad news? sentiment scoring of negativity, fear, and anger in news content,” *The ANNALS of the American Academy of Political and Social Science*, 659(1), pp. 108–121. Available at: <https://doi.org/10.1177/0002716215569217>.

[8] Tomz, M., Weeks, J.L.P. and Yarhi-Milo, K. (2019) “Public opinion and decisions about military force in democracies,” *International Organization*, 74(1), pp. 119–143. Available at: <https://doi.org/10.1017/s0020818319000341>.

[9] Venugopalan, M. and Gupta, D. (2015) “Exploring sentiment analysis on Twitter data,” *2015 Eighth International Conference on Contemporary Computing (IC3)* [Preprint]. Available at: <https://doi.org/10.1109/ic3.2015.7346686>.

[10] YASWANTH, V.A.N.A.M.A., MATHUR, V.I.D.H.I. and SINGH, S.N.E.H. (2022) “Sentiments analysis using tweets.” Available at: <https://doi.org/10.21203/rs.3.rs-1706745/v1>.