The Russian invasion of Ukraine through the lens of ex-Yugoslavian Twitter

Bojan Evkoski[1,2], Igor Mozetič[1], Petra Kralj Novak[3,1], and Nikola Ljubešić[1,4]

[1] Department of Knowledge Technologies, Jozef Stefan Institute, Ljubljana, Slovenia
bojan.evkoski@ijs.si
[2] Jozef Stefan Postgraduate School, Ljubljana, Slovenia
[3] Central European University, Vienna, Austria
[4] Faculty of Information and Communication Sciences, University of Ljubljana, Slovenia

The Russian invasion of Ukraine brings about dramatic changes in the world. Analysing the structure and content of the communication on social media, such as Twitter, can help in understanding the causes, developments and consequences of this conflict. The geographical area of interest in our research is a part of ex-Yugoslavia where the BCMS (Bosnian, Croatian, Montenegrin, Serbian) languages are spoken, official varieties of the pluricentric Serbo-Croatian macro-language. This area is strongly politically divided by diverging influences of NATO (Croatia, Montenegro, Northern Macedonia, Bosniak and Croatian entity in Bosnia and Herzegovina) and Russia (Serbia, Serbian entity in Bosnia and Herzegovina). While Croatia is full EU member since 2013, Montenegro, Northern Macedonia and Serbia are EU candidate members, while Bosnia and Herzegovina is a potential candidate. Regarding military alliances, NATO members are Croatia (since 2007), Montenegro (since 2017) and Northern Macedonia (since 2020), while Serbia does not aspire to join NATO, primarily due to a complex Serbia-NATO relationship caused by the NATO intervention in Yugoslavia in 1999.

To shed light on the impact of the Russian invasion on this brittle and complex geographical and political area, we use social network analysis over available Twitter data, 6-weeks before and 6-weeks during the invasion. We discover a complex landscape of ideology-specific and country-specific communities, and analyse the transition into evident pro-Ukraine and pro-Russia leanings. As the communities show very divergent properties, we echo concerns of the heavy polarization and possible destabilization of this area of the Balkans.

Results. The data analysed were collected with the TweetCat tool [2], focused on harvesting tweets of less frequent languages. TweetCat is continuously searching for new users tweeting in the language of interest by querying the Twitter Search API for the most frequent and unique words in that language. This data collection procedure is run for the BCMS set of languages since 2017. During the 12 weeks of our focus, we collected 1.2M tweets and 3.8M retweets from 45,336 users. A rough estimate of the per-country production of tweets via URL usage from country-specific top-level domains (upper part of Table 1) shows for Twitter to be much more popular in Serbia and Montenegro than in Croatia or Bosnia and Herzegovina.

We created **pre-invasion** and **invasion retweet networks** (users as nodes, retweets as edges) from the collected data. We applied community detection (Ensemble Louvain) on the two networks and analysed the community properties and user transitions [1]. We identified and named the large communities (more than 100 users) by careful analysis of their most influential users and hashtag/URL usage. Fig. 1 depicts the user transitions

between the two networks, while Table 1 shows general statistics of each community. We discovered the following peculiarities:

- The BCMS tweetosphere is dominated by Serbian (RS) users and content.
- The political communities are more active compared to the non-political ones.
- RS populist coalition community (led by Serbian president Aleksandar Vučić) forms a very strong echo chamber, with less than 2% of all users, yet more than 25% of tweets and retweets and more than 95% of intra-community retweets.
- RS populist coalition and left-wing opposition remain neutral on the invasion topic.
- RS right-wing opposition and the Bosnian Serbs show a clear pro-Russia stance.
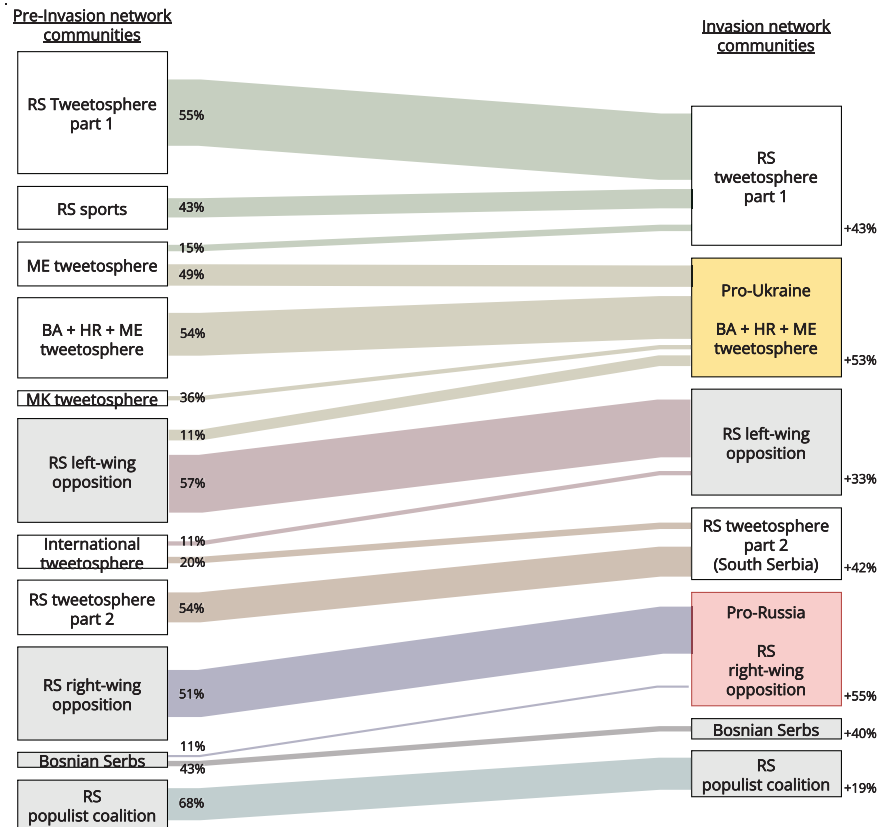- Croatian, Bosnian and Montenegrin communities show a clear pro-Ukraine stance.

Fig. 1. A Sankey diagram showing the transitions of users from the pre-invasion network communities (left) to the during invasion network communities (right). Rectangle height is proportional to the community sizes. Percentages near the pre-invasion communities show the portion of users found in the corresponding invasion communities. Percentages on the right-hand side of the invasion communities show the portion of users not previously present in the large communities of the pre-invasion network. Gray rectangles depict the communities tightly related to politics, with the yellow and red being the detected pro-Ukraine and pro-Russia leaning communities.

Future work. We continue with hashtag usage and media coverage analysis in the detected communities with focus on the polarization and ideological clash of BCMS

Country	Population	URLs
Serbia (RS)	**7.2M (47.3%)**	**106K (44.2%)**
Croatia (HR)	3.9M (25.6%)	19.6K (8.1%)
Bosnia and Herzegovina (BA)	3.5M (23.0%)	14.9K (6.2%)
Montenegro (ME)	620K (4.1%)	24.7K (10.2%)
Total	15.2M	242K

Pre-invasion communities	Users	Tweets	Retweets	Intra-com. RTs
RS tweetosphere part 1	**13K (29.0%)**	**125K (24.9%)**	300K (18.9%)	80.3%
RS tweetosphere part 2	2.5K (5.6%)	35.8K (7.1%)	63.2K (4.0%)	62.3%
RS sports	1.6K (3.6%)	12.6K (2.5%)	25.6K (1.6%)	53.8%
ME tweetosphere	1.7K (3.8%)	22.7K (4.5%)	44.6K (2.8%)	74.5%
BA + HR + ME tweetosphere	5.6K (12.4%)	37.8K (7.5%)	59K (3.7%)	75.3%
Macedonian tweetosphere	200 (0.4%)	721 (0.1%)	771 (0.1%)	77.7%
International tweetosphere	934 (2.0%)	8.5K (1.7%)	11.5K (0.7%)	62.3%
RS populist coalition	2.0K (4.8%)	52.4K (10.4%)	396K (24.9%)	**98.7%**
RS left-wing opposition	9.3K (20.6%)	105K (20.9%)	**408K (25.5%)**	80.5%
RS right-wing opposition	7.6K (16.8%)	87.8K (17.4%)	247K (15.5%)	72.1%
Bosnian Serbs	139 (0.3%)	2.2K (0.4%)	3.8K (0.2%)	83.1%
Total	45.3K	502.9K	1590K	

Invasion communities	Users	Tweets	Retweets	Intra-com. RTs
RS tweetosphere part 1	**16.9K (29.5%)**	160K (22.4%)	387K (16.8%)	71.1%
RS tweetosphere part 2	4.5K (7.7%)	57.3K (8.1%)	118K (5.1%)	58.1%
Pro-Ukraine BA + HR + ME tweetosphere	12.4K (21.7%)	76.1K (10.6%)	235K (10.2%)	64.7%
Pro-Russia RS right-wing opposition	11.1K (19.4%)	129K (17.9%)	508K (22.1%)	65.1%
RS populist coalition	1.8K (3.1%)	**208K (29.1%)**	450K (19.5%)	**95.6%**
RS left-wing opposition	9.8K (17.2%)	191K (26.7%)	**590K (25.6%)**	72.6%
Bosnian Serbs	356 (0.6%)	5.4K (0.7%)	7.1K (0.3%)	62.3%
Total	57.4K (+26.7%)	717K (+42.8%)	2302K (44.8%)	

Table 1. First part shows general population of each BCMS country and their respective tweet URL shares (.rs, .hr, .ba and .me). Second part shows the pre-invasion network communities with the number of users, tweets, retweets and intra-community retweets. The third part shows the same statistics for the invasion network communities. Grey rows depict political communities, while yellow and red show the pro-Ukraine and pro-Russia communities.

Twitter with respect to the conflict between NATO and Russia. Beyond obtaining interesting insights, we aim to explore two frequent issues in using social media for societal analyses: (1) uptake bias of specific social networks across countries and communities, and (2) entanglement of the main event with other large-scale events.

References

1. B. Evkoski, I. Mozetič, N. Ljubešić, and P. Kralj Novak. Community evolution in retweet networks. *PLoS ONE*, 16(9):e0256175, 2021.
2. N. Ljubešić, D. Fišer, and T. Erjavec. TweetCaT: a tool for building Twitter corpora of smaller languages. In *Proc. 9th Intl. Conf. on Language Resources and Evaluation*, pages 2279–2283, Reykjavik, Iceland, 2014. ELRA.