Context aware evolution of emoji sentiment reactions in a large Telegram community

Mints Oleksiy
SMK College of Applied Sciences
Klaipeda, Republic of Lithuania *oleksii.mints@smk.lt*

Abstract

This study examines how reaction emojis evolve beyond their traditional emotional associations within a large Russian-language Telegram community. Using lexicon-based sentiment scoring, temporal frequency charts, and cluster analysis, we analyzed 220,972 reacted comments from August 2021 to April 2025, focusing on four high-frequency reactions: 👍, 👎, ❤, and 🤡. Our findings revealed that emoji meaning is not constant and can change over time. Spikes in the use of certain emoji reactions coincide with periods of social turbulence, indicating them in real time and potentially enabling prediction. These findings expose community-driven semantic shift and demonstrate that reaction patterns provide weak supervision cues for identifying sentiment-context mismatches, which aids in moderation and crisis detection. The results are also important for training neural network models using community-annotated messages.

Keywords: emoji reactions, semantic shift, sentiment analysis, Telegram, weak supervision

1. Introduction

Sentiment analysis is important for ensuring safe and comfortable online communication. At its core, it allows us to quantify and interpret the emotional tone of messages from individual users and entire online communities of various sizes. In different communities, sentiment analysis can perform functions such as monitoring emotionally charged and polarizing content, identifying potential manipulation, moderating content, and monitoring participants' morale. Additionally, content monitoring on social media can serve as a source of weak signals for predicting crises of various scales [5].

Classical lexicon-based methods remain attractive because they are lightweight and language-agnostic, yet they assume that words and symbols carry stable, universal meanings. In practice this assumption falters in highly contextual arenas where users rely on emoji. Reaction emojis provide a crowdsourced layer of interpretation and have been used as weak labels for emotion classification based on collective approval or anger in [7]. Authors of [8] found that Facebook reaction frequencies serve as strong indicators of users' emotional attitudes. Further research revealed the significant role of Facebook reactions in politics and social issues [1].

Recent sentiment analysis research is characterized by the use of resource-intensive neural network methods that work with large amounts of data and an expanded range of modalities, including emojis. However, the primary research objective thus far has been to identify a fixed vector representation of message elements [3, 4]. At the same time, the author's own observations, supported by the recent publications [6], demonstrate the contextual plasticity of emojis in certain online communities. This is particularly true on platforms that allow users to use a large number of emojis, such as Telegram [6].

The insufficient study of the conditions and consequences of emoji contextual plasticity makes the following research questions relevant:

RQ1: How reliable are conventional sentiment assignments for reaction emojis in capturing their actual communicative roles in online communities?

RQ2: What additional functions can reaction emojis fulfill in online communities?

2. Methodology and Results

This study utilizes a mixed methods design combining large-scale quantitative text mining and qualitative analysis of significant cases. The study focuses on *@nlevshitstelegram*, a Russian-language Telegram channel with over 110,000 subscribers. The channel is primarily used by expatriates from Russia, Ukraine, and Belarus who live in Georgia. The research period spans from August 1, 2021, to April 1, 2025. This timeframe covers several significant political events in Georgia and will capture potential shifts in communicative norms.

Data collection routines were written in Python 3.12 using the Telethon API. Data preprocessing, statistical analysis, interactive exploration, and clustering were conducted in Orange Data Mining v. 3.38, using custom Python scripts for data preprocessing. The author can provide this dataset upon request.

Personal identifiers beyond the public display name were neither harvested nor stored. After deduplication and verification, the corpus comprised 448,502 messages, 422,423 of which were user comments. Of those comments, 220,972 carried at least one emoji reaction. The message with the most reactions received 1,198.

2.1. Data pre-processing

The Orange workflow section related to data preprocessing stage is shown in Fig.1.

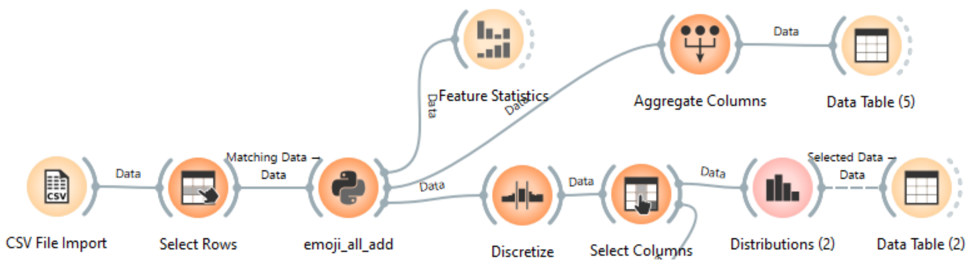

Fig. 1. Workflow of data preprocessing.

At this stage we removed all messages that were not user comments, split each reaction type into its own feature column, computed summary statistics for the sample, grouped the data by posting date to enable time-based analysis, and finally produced a histogram showing how many posts appeared in each period of the study (Fig. 2).

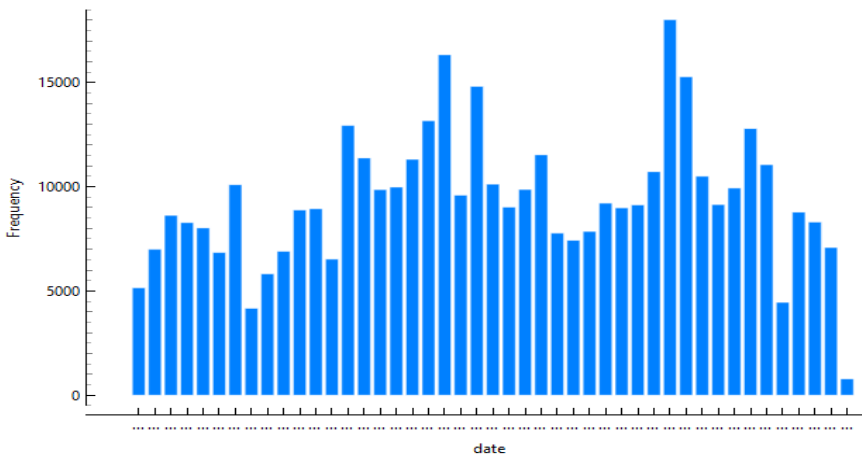

Fig. 2. Distribution of the message's frequencies.

Analysis of the distribution histogram shows that the average intensity of comments changed insignificantly throughout the entire period analyzed. Peaks are observed on the histogram, which correspond to the following periods: Mar 22; Sep 22; Mar 23; May 23; May 24; and Oct 24. Most of these periods are associated with political events, both internal and external, that strongly impacted Georgia and are reflected in the increased activity of the channel's subscribers.

2.2. Data analysis

Further, a statistical analysis of sentiments corresponding to the most popular emoji reactions was performed. During the analysis, only messages that raised emotional reactions from other users and were tagged with at least ten emoji were investigated. To evaluate sentiment in Russian-language messages, we used a unigrams-based multilingual sentiment lexicon described in [2]. During the research process, it became clear that a large proportion of the studied messages were evaluated as neutral (sentiment = 0). Therefore, statistics were additionally calculated for messages with a non-zero sentiment value. The results are shown in Table 1.

Table 1. Sentiment statistics for various emojis.

Emoji	total msgs with ≥ 10 emoji	mean (all)	median (all)	dispersion (all)	mean (excl 0)	median (excl 0)	dispersion (excl 0)
all	70410	1.51377	0	4.67294	2.76337	3	3.3926
👍	44120	1.56504	0	4.32875	3.03016	3.07765	2.28992
👎	7993	1.40581	0	5.2307	3.14332	3.125	3.64852
🤡	8450	1.52482	0	4.105	2.18176	2.74024	4.20452
❤	5952	1.96434	0	2.59	3.80136	2.94118	1.72805

Analysis of Table 1 shows that a significant proportion of emojis are weakly related to the emotional evaluation of text. Thus, the mean and median sentiment values for messages with opposite emojis, such as 👍 and 👎, are nearly identical. Additionally, these values are close to the corresponding values for the entire data sample. After analyzing individual messages and reactions to them, it was concluded that these emojis are primarily used to express agreement or disagreement with the message.

The data in Table 1 show that, among popular emojis, only ❤ can adequately indicate text sentiment. Its use correlates with higher sentiment values, so it can be considered an expression of gratitude for the message's content.

Now, let's look at the statistics for the 🤡 (clown) emoji. If we consider only messages with non-zero sentiment scores, the mean and median values for this emoji are much smaller than those for others. At the same time, the dispersion is larger. The frequency histogram of this emoji's use during different periods of the studied community's development (Fig. 3) is also very different from the others'.

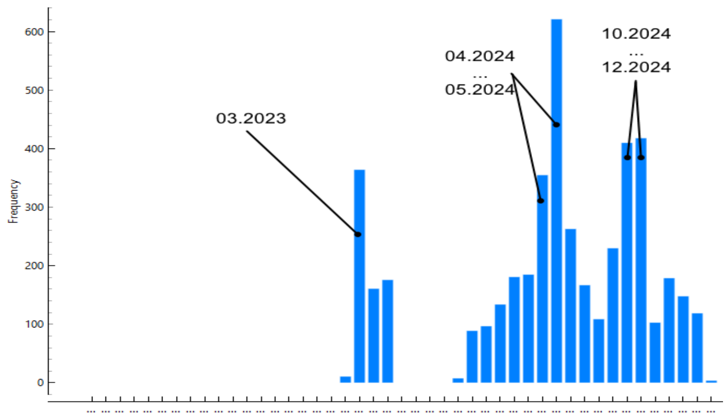

Fig. 3. Distribution of the 🤡 emoji frequencies.

Fig. 3 shows that the emoji "clown" became popular for the first time in March 2023. At that time, mass protests against the "Law on Transparency of Foreign Influence" began in Georgia. The law was removed from the agenda at the end of March 2023, and the protests ended. The next two peaks in Fig. 3 also coincide with political events in Georgia. From April to May of 2024, the government re-submitted the same law for consideration, and, despite the protests, passed it. Parliamentary elections were held in Georgia in October 2024, and the results caused mass protests.

During these events, there was increased activity by bots – program agents or people posting repetitive, manipulative messages in Telegram channels. These bots were difficult to detect automatically, but active participants in online communities learned to recognize them quite easily and started tagging them with the rarely used emoji 🤡. Over time, this practice spread to most of the channel's subscribers. Thus, in the community, this emoji began to fulfill additional functions, such as warning of bot activity. We can interpret this conclusion in the context of RQ2.

Also, as Fig.3 shows, from the end of 2023 emoji 🤡 began to be used more actively. Let's consider the peculiarities of its use with the help of text mining methods (Fig. 4).

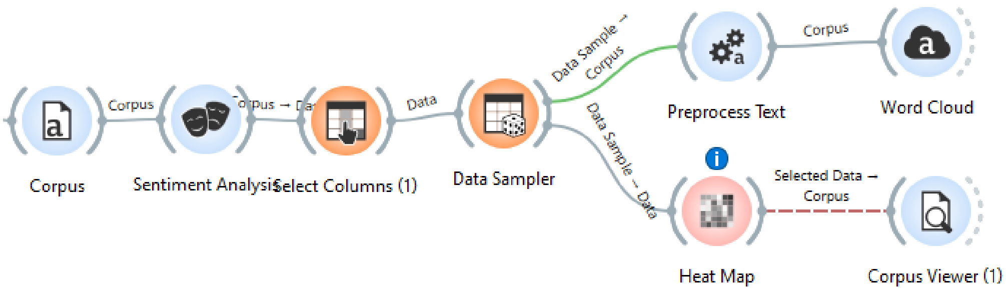

Fig. 4. Workflow of “clown” emoji analysis.

This workflow calculates sentiment predictions for messages containing the emoji 🤡 and clusters them using a heat map to identify the characteristics of different message groups. The analysis revealed the following:

- Among messages with high sentiment scores, many can be regarded as sarcastic. In this case, emoji 🤡 is an indicator of sentiment-context mismatches.

- Among messages in the same cluster, there are often nearly identical messages. This confirms the observation of high bot activity in the studied community.

A Word Cloud analysis of messages tagged with this emoji revealed a significant number of words with political connotations. These words are among the most frequent nouns and appear 2-3 times more frequently than in the entire sample (Table 2).

Table 2. Statistics on using politics-related words.

Word	In a subset of the general dataset	In the messages, marked by “clown”	Ratio
Грузия (Georgia)	1093	2218	2,03
Россия (Russia)	269	841	3,13
Украина (Ukraine)	<100	231	>2,31
грузин (Georgian)	271	652	2,41
страна (country)	433	1357	3,13

These calculations show that the 🤡 emoji is a highly likely sign of polarizing content in the analyzed data that causes conflicts among community members. During the period under review, such content was most often political.

Thus, in the context of RQ1 we can conclude that the communicative role of the most frequently used emoji reactions has shifted in the online community under study.

3. Concluding Discussion

This study contributes to the rapidly growing corpus of studies on multimodal sentiment analysis. Unlike studies [3, 4], which use machine-assigned fixed vector representations for emojis, our study empirically reveals how users redefine these symbols. Answering RQ1, our analysis shows that the "universal" emoji evaluations embedded in most automated tools are ineffective within specific communities. Some emojis lose their connection to the emotional context over time, while others acquire a metacommunicative function.

The most obvious contextual shift observed in the study was the transformation of the 🤡 emoji response from an ironic symbol to a warning about bot activity and polarizing content. Changes in the meaning of other common symbols, such as 👍 and 👎, were also observed. In contrast, the meaning of the ❤ response did not change and remained positive. Most likely, the shift in the communicative role of emoji reactions is characteristic of platforms that provide users with a large choice of them when communicating, such as Telegram or Discord.

Another important finding was obtained by answering RQ2. We have shown that emoji reactions are a form of message annotation in self-organized online communities. This has practical implications for crowd-sourced moderation systems. Also, labelling messages with emoji reactions is important for developing AI sentiment analysis models, as it allows evaluation of the author's words through a collective sentiment filter.

Acknowledgements

The author would like to thank his son Mints Mykhailo for assistance in preparing the program for data collection.

References

1. Anwar, S., Giglietto, F.: Facebook reactions in the context of politics and social issues: a systematic literature review. Front. Sociol. 9, 1379265 (2024)
2. Chen, Y., Skiena, S.: Building sentiment lexicons for all major languages. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Vol. 2 (Short Papers), pp. 383–389. ACL (2014)
3. Kusal, S., Patil, S., Kotecha, K.: Multimodal text-emoji fusion using deep neural networks for text-based emotion detection in online communication. J. Big Data 12 (1), 1–25 (2025)
4. Lou, Y., Zhou, J., Zhou, J., Ji, D., Zhang, Q.: Emoji multimodal microblog sentiment analysis based on mutual attention mechanism. Sci. Rep. 14 (1), 29314 (2024)
5. Mints, O., Simanavičienė, Z., Polupanova, K., Mavrina, M.: A Conceptual Framework for Early Diagnosis of Global Catastrophes Based on Information Field Monitoring. In: Proceedings of the 47th Eurasia Business and Economics Society Conference, Eurasian Business and Economics Perspectives, vol. 34, p. 299. Springer Nature (2025)
6. Morales, E., Hodson, J., Gruzd, A., Mai, P.: I👍 your Hate: Emojis as Infrastructural Platform Violence on Telegram. In Proceedings of the 58th Hawaii International Conference on System Sciences, pp. 2316-2325 (2025)
7. Pool, C., Nissim, M.: Distant supervision for emotion detection using Facebook reactions. arXiv:1611.02988 (2016)
8. Tian, Y., Galery, T., Dulcinati, G., Molimpakis, E., Sun, C.: Facebook sentiment: reactions and Emojis. In: Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media, Valencia, Spain, Association for Computational Linguistics, pp. 11–16 (2017)