Emoji Trends Between Abnormal Usage and Cultural Differences: A Case Study of Emojis

Ahmed Almohanadi
Tokyo Metropolitan University
Tokyo, Japan
email:ahmed-almahanadi@ed.tmu.ac.jp

Shohei Yokoyama
Tokyo Metropolitan University
National Institute of Informatics
The University of Tokyo
Tokyo, Japan
email:shohei@tmu.ac.jp

***Abstract*—Emojis can be used to understand people's emotions through non-lexical language around the world, and through different techniques, they can be used to study and understand people's trends and cultural traits and gauge abnormal activities or events. This paper analyzes a monthly data period consisting of multiple activities where unique trends and abnormal emojis are highly likely to occur. An established method has been created, and map projection was used to establish this relation. We found interesting cultural traits, special events, and coronavirus pandemic-related emojis projected on a clustered map, showing a strong relation between certain emojis and specific regions. All these are common emojis that can be seen daily and unique emojis that appear in special events.**

***Keywords*—Emoji; Twitter; Clustering; Geolocation Data; Data Analysis.**

I. Introduction

Emojis in today's online communication are very popular due to their brevity, conciseness, and ability to give the message an emotional coloring. Moreover, these special characters greatly facilitate global communication, allowing people to convey ideas and even complex narratives concisely and understandably.

As a result, emojis have become an integral part of almost any digital communication, from interpersonal communication to brands and companies communicating with their customers. Since emojis have become commonplace in modern communication, trends in their use have been repeatedly explored.

According to Kaur et al. [1], emojis can be used to portray emotional reactions among social media users who understand the emoji similarly. The images facilitate non-verbal communication cues on social media that enhance user exchange and interaction.

Using Emojis is a dynamic concept. The Emoji set in smartphones is constantly updated as a reaction to events in the world. Moreover, often, long-standing images take on new meanings, reflecting the ever-changing landscape of cultural, social, and technological development.

This evolution provides researchers with a unique opportunity to delve into the complex relationships between emojis and the societies that use them. In addition, emojis contribute to formulating and expressing public opinion regarding significant incidents and events.

Therefore, this research paper analyzes the emoji trends that are common today, as well as the features of their formation due to geographic and cultural differences between users. Analyzing the patterns and frequency of emoji usage across countries and regions, this study seeks to discover an exciting insight into the cultural nuances of digital expression. Since emojis are a cultural characteristic, this study suggests that different figures' usage patterns and meanings may differ from country to country.

Moreover, countries and regions may have localized symbols reflecting the interplay between regional cultures and digital language, providing a glimpse into the nature of online communication worldwide. An analysis of the geographical features of Emoji will help establish the universality of this form of communication and try to understand the global trends.

The prominence of emojis applies across many parts of the world, from the West to the East. According to Dyer and Kolic [2], the state of California in the United States has been found to have a high concentration and multiplicity of emoji usage, especially along the coastal area. The trend is connected to the high population in the state's central coastal cities, including San Diego, Los Angeles, Palo Alto, San Francisco, and others (Kejriwal et al., [3]). The high usage is also correlated to the linguistic and cultural diversity of the area, which promotes greater exchange and interaction.

The rest of this paper is organized as follows. Section 2 reports related work. Section 3 describes our proposed method. Section 4 presents the results and findings in three subsections. Section 5 concludes the paper and discusses future work.

II. Related Work

The peculiarities of using emojis have attracted considerable attention from researchers. However, due to geographical factors, the distinctive features of this type of communication have been identified relatively recently.

Moreover, most researchers in this field emphasize more than the influence of the region's geographical features but the cultural trends of the countries that determine the use of Emoji [4]. Assessing the influence of the country's cultural

Copyright (c) IARIA, 2023. ISBN: 978-1-68558-103-9

characteristics on the use of Emoji, the researchers found variability in the interpretation of Emoji in different cultures.

Moreover, the fewer ways of communication between different countries and cultures, the more pronounced the difference in the use and meanings of emojis. For example, Guntuku et al. [4] found that there are significant differences in the interpretation and frequency of emoji usage between the East (China and Japan) and the West (United States and United Kingdom). Representatives of Eastern cultures often use text messages instead of using the corresponding figures.

In addition, since emojis were initially conceived as a way to color messages emotionally, their meaning in different cultures is influenced by general trends regarding the manifestation of emotions and feelings. Continuing to analyze the difference in emoji usage between Western and Eastern countries, Gao and VanderLaan [5] found that Westerners and Easterners appear to perceive emotions differently.

This difference in perception has been established even with basic emotions like happiness and sadness. Therefore, the researchers concluded that people might perceive the intent of the communication more accurately during exchanges with others of the same, compared with a different cultural background. This casts doubt on the claim that Emojis are a universal way of global communication.

On the other hand, using these figurines has demonstrated its ability to facilitate intercultural communication. Research has explored how these symbols bridge language barriers and facilitate interaction between different language communities, enriching global digital discourse.

In addition, the development of Emoji contributes to forming a diversified and inclusive world society. For example, the appearance of images with different skin colors or pairs in different combinations contributed to the unification of people since everyone can express themselves.

Despite the presence of regional differences, Kimura-Thollander and Kumar [6] note that the use of Emoji contributes to global cultural exchange and the formation of a common background among representatives of different communities.

Moreover, the participants in their survey came to the conclusion that it is necessary to create universal emojis that will have the same meaning in all countries of the world. Such an approach has the potential to simplify global communication but, at the same time, threatens to destroy global cultural diversification.

The use of emojis is constantly evolving and adapting to changing human needs. Therefore, many people attach particular meanings to individual images that are understandable only to them. Schouteten et al. note that in modern digital interpersonal communication, it is common for people to use emojis as a kind of shorthand.

At the same time, this form of communication may be meaningless for other people. Brands often use the same approach to create a specific image of the company in the consumer's perception. Distributing unique, localized emojis works the same way, with a group of people defined according to their nationality or where they live.

After collecting statistics on the most popular emojis in different countries, Cohen [7] determined that the red heart and the laughing face with tears are the most popular around the world. At the same time, the national flags of countries are in second place in popularity, which expresses their patriotic sentiments.

A rather remarkable figure is the image of two folded hands, which in some cultures is interpreted as a high-five, while in others, such as India, it expresses gratitude [8]. These patterns are due to distinctive cultural, linguistic, and national features that are characteristic of a particular region.

Finally, Emojis are a universal way to express public sentiment on social media. In addition, emoji usage trends may change in response to national situations or significant global events. Analyzing user pages on Twitter allows you to determine public opinion regarding certain global events, for example, as was the case with the COVID-19 pandemic. At the same time, due to the global nature of the incident, corona-related emojis are universal and have the same meaning regardless of region or country. In addition, there are unique emojis that do not have global popularity but are often used by certain groups of people. Examples of such images are sunflower, unicorn, rocket, artist palette, and top hat emojis [9].

These symbols may not be universally accepted, but their meaning in a particular community makes them invaluable for conveying subtle meanings. For example, they can be a manifestation of cultural self-identification, belonging to a particular subculture, and a form of self-expression. Unique emojis often have a distinct visual style that draws attention and catches the eye, which is why they are gaining popularity among social media users.

III. METHOD

In this study, the data collected for the research comprised using Twitter analysis tool for the period selected was one month between the 19th of November and the 18th of December of 2022; during this period, the Qatar World Cup 2022 was conducted, and Winter started in many areas of the world.

The gathered tweets were filtered with two criteria: first, they must include geolocation information, and second, the tweet must be classified as English according to the Twitter API dataset.

In the next step, the data was processed through multiple functions to be ready for final usage, including filtering from spam and duplication. The geolocation data was converted into radians to calculate the distance between points and establish density-based clusters.

For the abnormal analysis method, all emojis were extracted from the text each day of the month and used to calculate the Term Frequency–Inverse Document Frequency (TF-IDF) method, which is used to estimate the frequency and importance and calculate the z-score for each Emoji for the perspective day. The z-score threshold was set, and a search

Copyright (c) IARIA, 2023. ISBN: 978-1-68558-103-9

for a high difference in this value was conducted to search for abnormality.

Lastly, the map clustering method used was Density-based spatial clustering of applications with noise (DBSCAN).

IV. Result

In this research paper, we have divided the results into three parts based on the clustering type and result category we have found.

A. Corona Related Emoji Map Clustering

In this experiment, random emojis were fed to the clustering map to find a unique emoji with a massive presence in one area besides the others. The corona-related Emoji, the mask face (😷), was populated within the US area, which matched a period where the coronavirus spread in another wave.

B. Culture Related Emoji Map Clustering

Another interesting Emoji that popped out was the skull emoji (💀). Its population was focused on African countries on the west coast of Africa. After studying the Emoji, reading the comments, and further research, it is inferred that the meaning of skull emoji has the same meaning as the Emoji with the Face with Tears of Joy Emoji (😂). It is inferred that dark humor in the Emoji type is used to indicate interest, awkwardness, and irony.

C. Abnormal Emoji Trend

In this experiment, TF-IDF was used to search for abnormal usage of all emojis; using a z-score table, the maximum threshold estimated for the period was 4, a very high value that resulted in many low spike points due to having many emojis being used extreme rare cases, most of them are random usage. Three interesting findings were made where an unexpected high point was estimated.

a) Nazar (amulet) emoji 🧿*:* This emoji had a peak on the 30th of 11 It was found mostly in America and India shown in Fig 1 It is a special day related to religion, marriage, and personal matters across all regions; many people are getting married on this day. The combination of religious Indian culture, with an interesting number of days, along with personal matters, made this Emoji spike.

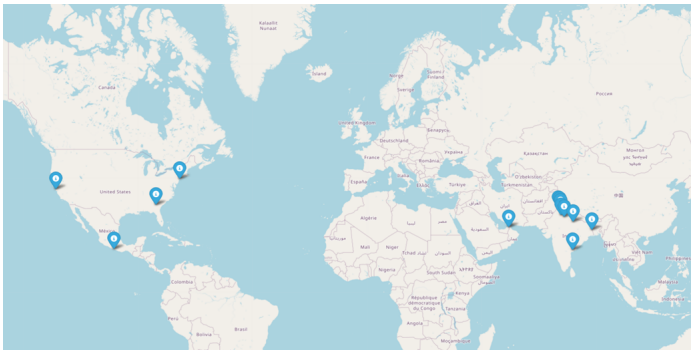

Fig. 1. Clusters mostly found in India and the United States

b) Goat emoji 🐐*:* This Emoji has been recently referred to by its acronym GOAT as Greatest Of All Time, which is typically used for superstars across all fields, including music, art, gaming, and sports. It spiked on the 14th of December, shown in Fig 2, because of the awaiting match between Argentina and Croatia, where Argentina won 3-0 with Lionel Messi opening the scoring and getting the MVP for that match. Messi has been called goat by fans across the world the GOAT, considering him the best player in the world, and he did win the 2022 Qatar World Cup with his team.

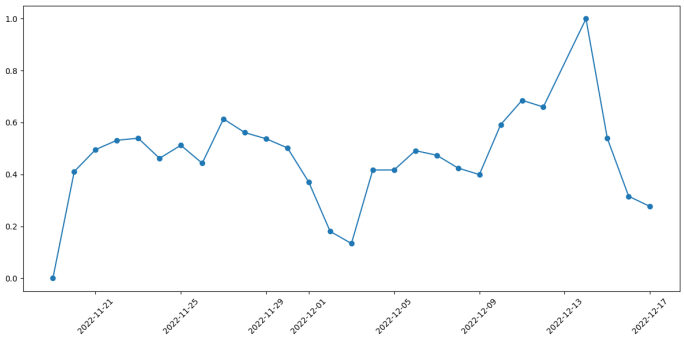

Fig. 2. Sentimental Values of the collected sample

c) Weather emojis 🌀 *and* 🌧️ *:* Both of the emojis have experienced a spike on the 14th of December. Going by the map clustering, the most significant spikes are in two areas: the east coast of the US and the second one is the UK and around the England channel areas. It was reported on the 14th of December that thunderstorms were spotted from New York on the east coast; however, in the UK, it was snowing beautifully, and it was one of the most beautiful days to share on social media. Therefore, the simultaneous events caused the spike in these emojis together.

V. Conclusion and Future Work

Clustering tweets with Geolocation and Emojis have yielded exciting results, from pandemic conditions and weather confirmation to unique traits and global trends. However, this study has been limited to many factors that might have reduced or hindered exciting discoveries.

The selected period was interesting for many reasons, and events happened all at the same time of that year. However, other periods may provide different results that can be interrupted differently.

The data sample was filtered using English language tweets where it is spoken globally. This restriction was necessary because local languages created massive clusters in their countries, causing an imbalance in the results. With that being said, it might yield interesting results when including different languages together, and it will definitely affect the result of the culture-related Emoji map.

An overall point to point out is that people with high-security alerts toward privacy tend to limit access to their location and prevent social media from accessing any related information. Such trivial denial of access can cause such kinds

Copyright (c) IARIA, 2023. ISBN: 978-1-68558-103-9

of studies to lose their opportunity to add value to an important field that could change understanding in NLP studies.

We urge to spread awareness and ask the general people to allow providing location access and show the importance of the studies being done on such data and the current protection provided by the companies where location is not provided accurately for other parties but somewhat slightly adjusted for the provider's safety.

TF-IDF has generated reliable examples, but we were expecting more results; therefore, an investigation for a different abnormal spike detection should be prioritized. Detecting by number count or trend might require training a model to understand the different emojis and types of spikes required for the study.

DBSCAN has generated reliable clusters that could be analyzed and examined. Although other methods were considered, they were deemed inferior due to the outstanding performance of DBSCAN in these kinds of studies. However, this should be reconsidered as new methods are getting updated rapidly.

Overall, if added, automating the process within a long-term period can yield more interesting findings; the next phase will focus on automation and training models that can detect and provide the most interesting results in terms of clustered areas and abnormal changes, discovering unique traits unseen discoveries in the emoji language that is being used on a daily basis.

ACKNOWLEDGMENT

This research was conducted as part of "COVID-19 AI & Simulation Project" run by Mitsubishi Research Institute commissioned by Cabinet Secretariat, JAPAN

REFERENCES

[1] Kaur, S., Kaul, P., & Zadeh, P. M. (2020). Monitoring the dynamics of emotions during COVID-19 using twitter data. Procedia Computer Science, 177, 423–430.

[2] Dyer, J., & Kolic, B. (2020). Public risk perception and emotion on Twitter during the Covid-19 pandemic. Applied Network Science, 5(1).

[3] Kejriwal, M., Wang, Q., Li, H., & Wang, L. (2021). An empirical study of emoji usage on Twitter in linguistic and national contexts. Online Social Networks and Media, 24,

[4] S. C. Guntuku, M. Li, L. Tay, and L. H. Ungar, "Studying cultural differences in emoji usage across the east and the west," Proceedings of the International AAAI Conference on Web and Social Media, vol. 13, no. 1, pp. 226-235, 2019. [Online]. Available: https://doi.org/10.1609/icwsm.v13i01.3224. [Accessed: the 16th of September, 2023].

[5] B. Gao and D. P. VanderLaan, "Cultural influences on perceptions of emotions depicted in emojis," Cyberpsychology, Behavior, and Social Networking, vol. 23, no. 8, pp. 567-570, 2020. [Online]. Available: https://doi.org/10.1089/cyber.2020.0024. [Accessed: the 16th of September, 2023].

[6] P. Kimura-Thollander and N. Kumar, "Examining the 'global' language of emojis: Designing for cultural representation," Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 495-508, 2019. [Online]. Available: https://doi.org/10.1145/3290605.3300725. [Accessed: the 16th of September, 2023]. [Accessed: the 16th of September, 2023].

[7] B. Cohen, "The most popular emoji in each country around the world," The Gate, the 15th of May, 2022. [Online]. Available: https://thegatewithbriancohen.com/the-most-popular-emoji-in-each-country-around-the-world/ [Accessed: the 16th of September, 2023].

[8] M. Kejriwal, Q. Wang, H. Li, and L. Wang, "An empirical study of emoji usage on Twitter in linguistic and national contexts," Online Social Networks and Media, vol. 24, pp. 1-23, 2021. [Online]. Available: https://doi.org/10.1016/j.osnem.2021.100149.

[9] J. J. Schouteten, F. Llobell, S. L. Chheang, D. Jin, and S. R. Jaeger, "Emoji meanings (pleasure–arousal–dominance dimensions) in consumer research: Between-country and interpersonal differences," Journal of Food Science, vol. 88, no. 1, pp. 106-121, 2023. [Online]. Available: https://doi.org/10.1111/1750-3841.16374. [Accessed: the 16th of September, 2023].

Copyright (c) IARIA, 2023. ISBN: 978-1-68558-103-9