The World Wants Mangoes and Kangaroos: A Study of New Emoji Requests Based on Thirty Million Tweets

Yunhe Feng[1], Wenjun Zhou[1], Zheng Lu[1], Zhibo Wang[2], Qing Cao[1]

[1]University of Tennessee, [2]Wuhan University

Motivation

- As the usage of emojis (and social media in general) evolves, new emojis are being continuously requested.
- The new emojis proposal submitted to the Unicode Consortium completely ignores emoji petitions directly generated by users who have first-hand information regarding the valuable usage context.
- A systematic study on which new emojis are wanted, when, where and why these emojis are requested, and how to call for these emojis still remains unexplored.

Data Curation

- More than thirty million tweets containing the word "emoji" were collected from Oct. 2017 to Oct. 2018 using Twitter's Streaming APIs.
- We filtered out bot-generated tweets by removing eleven users who produced on average more than 10 "emoji" tagged tweets per day and 131 users who tended to tweet regularly every few minutes.
- A series of preprocessing steps were applied to reduce the frequent presence of abbreviations, irregular expressions, ill-formed words and non-dictionary terms.
- We extracted user profiles, tweet contents, timestamps, and geo tags, from tweet JSON files, and built the *complete dataset* (general tweets, retweets, quoted tweets and replies) and the *unique dataset* (only original tweets) for further exploration.

Emoji Extraction Using Linguistic Patterns

We summarized 49 frequent linguistic patterns and their 2620 variations to match emoji-requested tweets and extract emojis. More linguistic patterns and corresponding tweet screenshots are available through http://yunhefeng.me/linguistic_patterns.html.

- why is there no *foo* emoji
- where is the *foo* emoji
- need a *foo* emoji
- look for a *foo* emoji
- demand a *foo* emoji
- have no *foo* emoji
- invent a *foo* emoji
- a *foo* emoji is overdue
- still no *foo* emoji

Requesting Extant Emojis

When examining extracted emojis, to our surprise, we found hundreds of emojis that had already been released by the Unicode Consortium were still requested extensively. Figure 1 demonstrates the top 25 most requested extant emojis. Seven out of the top 10 extant emojis came from the recent Emoji Version 5.0, which was released in May 2017, while we started to crawl the data in Oct. 2017. It is also interesting to note that the percentage of Twitter users on mobile was about 80%, but they contributed more than 91.8% extant emoji requests, as shown in Figure 2.

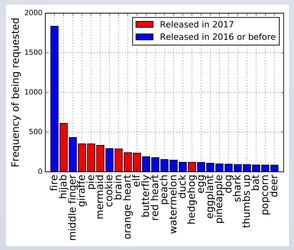

Fig. 1: Top 25 most requested extant emojis

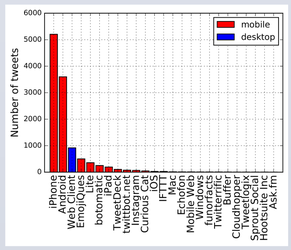

Fig. 2: Platform distributions

Requesting New Emojis by Category

We proposed a WordNet based emoji classifier to cluster new emojis requested by users into the following eight categories. More than 31.8% of wanted emojis were from the Smileys & People category, which might indicate people's great passions for new emojis to express emotions. The public also desired many emojis, including kangaroos and mangoes, from the Animals & Nature and Food & Drink categories.

Category	# emojis	# tweets	examples
Smileys & People	385	36,790	redhead, ass shaking
Animals & Nature	185	18,059	kangaroo, flamingo
Food & Drink	164	12,067	mango, waffle
Activity	42	2,421	slide, softball
Travel & Places	56	1,946	compass, brick
Objects	161	12,229	broom, red carpet
Symbols	170	21,443	anarchy, infinity
Flags	44	3,156	trans flag, Texas flag

Temporal Distributions of Wanted Emojis

Figure 3 demonstrates emojis that were requested more than 1000 times throughout one year (Oct. 2017 - Oct. 2018). Although the overall requested number was not very large, emojis of brooms, flamingos and kangaroos appeared consistently in all months. In contrast, heavily requested emojis like the lookout and the red carpet mainly appeared in one or two months. The fact that the broom, flamingo and kangaroo emojis were selected as part of Unicode 11.0 in 2018 or Unicode 12.0 in 2019 implied those emojis that were requested continuously and by multiple users were more likely to be approved by the Unicode Consortium as they reflected the real needs of the majority of online users. We also found the extensive but relatively concentrated emoji requests were usually triggered by celebrities or their followers.

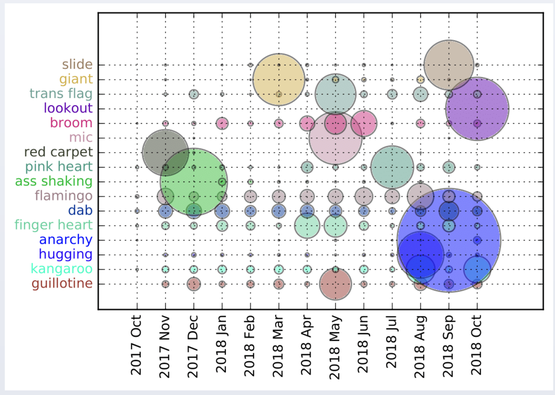

Fig. 3: The number of requests per emoji per month

Geographic Distributions of Wanted Emojis

English-speaking countries, such as United States (73.6%), United Kingdom (10.9%), and Canada (3.2%), contributed the most of emoji requests. States like California, Texas, and New York made a large number of requests. After normalizing by state population, the geographical distribution was relatively smooth and even across the country, which indicated that people in different states had a similar level of desire for new emojis.

Fig. 4: Requests per state in the U.S.

Fig. 5: Requests normalized by state population

Context of Emoji Requests

- Time-related events & activities
  - Time-sensitive and content-related emojis: candy cane emoji on Candy Holidays; the carnation emoji on Mother's Day; the waffle emoji on National Waffle Day
  - Popular entertainment products or events: Black Panther emojis; Olympic Rings emoji in 2018 Winter Olympics; yellow/red card emojis in 2018 FIFA World Cup
- Place-related interests
  - A single landmark: an Eiffel Tower emoji was petitioned in Paris
  - Tourist attractions: Mickey & Minnie emojis were requested at Walt Disney World
  - Regions or Countries: residents in Hawaii and Texas looked for state flag emojis
- Twitter influencer-related behaviors
  - Emoji requests made by prominent people on Twitter might trigger a widespread discussion of the requested emojis through a massive number of followers.

Advocacy Behaviors

When wanted emojis were unavailable, nearly one in three Twitter users would use the symbol of '@' to mention some people or organizations for their attention. More than 12% of users inserted #hashtags in their tweets when wanted emojis were inaccessible.

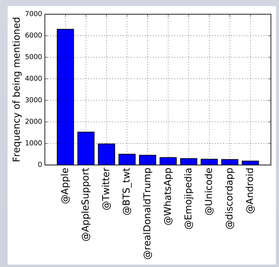

Fig. 6: @Mentions

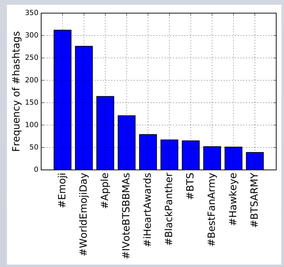

Fig. 7: #Hashtags

Relatedness, Fairness, and Equality

An interesting scenario for the emoji request rose when users complained that there existed an emoji for A but no emoji for B in the same tweets. They thought it was unfair or unreasonable because A and B were usually very similar or related to each other.

- Gender, color, similar function and looking cause a sense of unfairness and inequality.
- Emerging technologies, recent social movements, and the equality of political symbols motivate people to petition for new emojis.

Domain	Available Emoji (A)	Unavailable Emoji (B)
Human Diversity	breast-feeding	male-holding-baby
	man in tuxedo	woman in tuxedo
	blond hair	red hair
Life	pancakes	waffle
	bed	pillow
	wine glass	white wine
Science & Tech	antenna bars	Wi-Fi
	microscope	DNA
	mobile phone	ipad/tablet
Nature	honeybee	fly
	tortoise	turtle
	crab	lobster
Business	Unicode (U+20BF)	Bitcoin
	TOP arrow	bottom arrow
	bar chart	pie chart
Society	#MeToo hashtag	#MeToo in Unicode
	Greenland flag	transgender flag
	water pistol	real gun (AR15)
Politics	elephant for GOP	donkey for Dems.
	Guyana/Ghana flag	pan African flag
	United States flag	Confederate flag
Entmt. & Arts	trophy	Oscar
	videocassette/DVD	cassette tape
	guitar	ukulele

Significance of New Emojis

- Petitions for unreleased emojis

The newly added hijab (woman with headscarf) emoji through the Hijab Emoji Project campaign promotes inclusivity for about 550 million Muslim women on this earth.

Researchers proposed a mosquito emoji to better explain mosquito-borne illnesses like malaria, Zika, dengue and yellow fever.

- Petitions for new appearances of extant emojis
  - Twitter's lobster emojis were missing a set of legs
  - Misplaced cheese in Google's burger emoji
  - Lackluster appearance of Apple's bagel emoji
  - Curb visual representations of gun violence (Apple, Google, Microsoft, Facebook, Twitter)

Conclusion

- We revealed new and strong evidence of frequency, i.e., the explicit and accurate evidence like "why is there no foo emoji", for the Unicode emoji community to evaluate emoji petitions.
- Our study explained why extant emojis were still requested by users, and provided multiple suggestions to enable the timely availability of newly released emojis to users.
- A comprehensive understanding of new emoji requests was offered by profiling spatiotemporal distributions, summarizing advocacy behaviors, and exploring factors that inspire requests.
- We discussed the equality, fairness and diversity in emojis, and presented the potential significance of new emojis in many aspects like business promotion and violence control

Yunhe is looking for postdoctoral positions. More information about him is available via www.yunhefeng.me or the QR code.

Yunhe is looking for postdoctoral positions. More information about him is available via www.yunhefeng.me or the QR code.