Poster: Bringing Context into Emoji Recommendations

Joon-Gyum Kim, Taesik Gong, Evey Huang, Juho Kim, Sung-Ju Lee
KAIST

Bogoan Kim, JaeYeon Park, Woojeong Kim, Kyungsik Han, JeongGil Ko
Ajou University

ABSTRACT

We present Reeboc that combines machine learning and k-means clustering to analyze the conversation of a chat, extract different emotions or topics of the conversation, and recommend emojis that represent various contexts to the user. Instead of simply analyzing a single input sentence, we consider recent sentences exchanged in a conversation. we performed a user study with 17 participants in 8 groups in a realistic mobile chat environment. Participants spent the least amount of time in identifying and selecting the emojis of their choice with Reeboc (38% faster than without emoji recommendation).

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Text input**; *Smartphones*; • **Computing methodologies** → **Machine learning approaches**.

KEYWORDS

Emoji recommendation; mobile applications; machine learning; user experience

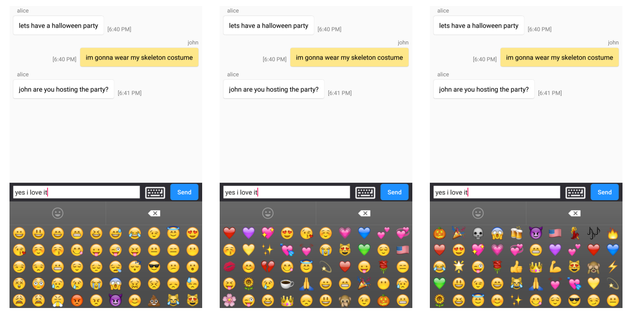
1 INTRODUCTION

As emojis are increasingly used in everyday online communication such as messaging, email, and social networks, various techniques have attempted to improve the user experience in communicating emotions and information through emojis. Emoji recommendation is one such example in which machine learning and natural language processing is applied to predict which emojis the user is about to select, based on the user's current input message. While emoji suggestion helps users identify and select the right emoji among plethora of emojis, analyzing only a single sentence leads to recommending many emojis of similar sentiment or emotion and missing various emotions or contexts expressed in conversations. Moreover, as the models analyze the current input sentence, they cannot suggest emojis to "emoji only sentences." As our user study indicates, 36.2% of the emoji-used messages were emoji-only inputs without any text. Existing recommendation models turn back to the default emoji layout for the emoji-only sentences as they have no input text to analyze.

We propose *Reeboc* (Recommending emojis based on context) that recommends emojis based on conversation context. Our study demonstrates the effectiveness of using beyond one chat line in emoji recommendations. Figure 1 shows an example of emoji recommendation during mobile chat. The default keyboard (Fig. 1 (a)) displays the same layout whenever a user wants to send an emoji. An emoji recommendation model that considers the current sentence (Fig. 1 (b)) shows a different layout, focusing on the word "love" of the sentence. *Reeboc* (Fig. 1 (c)) instead analyzes the conversation beyond the current sentence and displays recommendations that capture various contexts appeared in the chat.

MobiSys '19, June 17–21, 2019, Seoul, Republic of Korea

ACM ISBN 978-1-4503-6661-8/19/06.
https://doi.org/10.1145/3307334.3328601



(a) The default layout. (b) The single-sentence model. (c) The conversation model.

Figure 1: Emoji suggestion example.

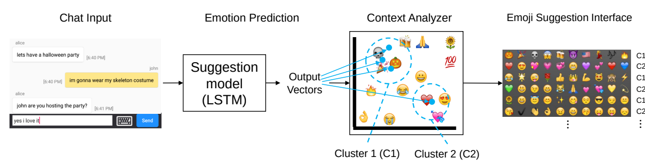


Figure 2: Emoji suggestion process.

2 EMOJI RECOMMENDATION BASED ON CONTEXT

We design our emoji recommendation system with the following considerations. First, our model must consider conversations, instead of using a single sentence for emoji prediction. Conversations not only include the text (and emojis) exchanged between the participants, but also the identity of the person who typed the sentence (i.e., the speaker) and the sequence of a chat sentence. Second, our model must capture various contexts of a conversation. There could be many different emotions, information, and sentiments expressed through conversations. The selection of emojis could be based on various contexts in an ongoing chat, and the recommendation model should provide the users with emojis that represent various contexts. Third, our model should provide recommendations for emoji-only sentences without any text input. By analyzing the conversation as a whole, not just the current input text, our model suggests relevant emojis to the users. Fourth, recommendations must be made in real-time.

We crawled our training and testing data from Twitter, specifically targeting conversations through the reply function on Twitter. We collected a total of 6.3 million English tweets from 1.5 million conversations from 1.8 million unique users.

We convert each word and emoji to a vector-form using word-to-vector [2]. In our system, we use the pre-trained Natural Language Processing Group (TALN) Word2Vec model [1], which has mappings from a word (word-to-vector) and an emoji (emoji-to-vector) to a 300 dimensional vector, respectively. Through the word-to-vector and emoji-to-vector conversions, we represent words and emojis as 300-dimensional vectors for making input features of LSTM.

To suggest appropriate emojis according to the chat context, an emoji suggestion system should take a number of previous messages (i.e., a conversation) into account. Our system feeds the last five messages and then provide five inputs into the LSTM model. We empirically decided on five samples as it helps us detect multiple emotions on a conversation to capture the context low latency. The LSTM model then generates five different output vectors and each would provide a meaningful emotion, possibly different from each other. These emotions and/or information will be represented as different clusters.

We cluster the output vectors of similar contexts through K-means clustering. The distance between the vectors is calculated as their cosine similarity. Our system finds that the output vectors of all clusters have the minimum number of clusters with a cosine similarity of at least 0.9 from the centroid, while controlling the K value of K-means clustering from 1 to the number of output emoji vectors. A centroid of the cluster is the mean value of the unit vector of each cluster. We empirically selected the cosine similarity criterion as 0.9 from our twitter data.

Reeboc provides recommended emojis based on the context clusters. Each cluster has its emoji list that is sorted by the cosine similarity between the centroid and emoji vectors in descending order. With these emoji lists from the clusters, *Reeboc* application (the rightmost image in Figure 2) displays the suggested emojis when a user clicks the emoji button. Each row of the "emoji pad" represents the sorted emoji list in each cluster (i.e., each detected context, emotion, or information).

3 USER STUDY

We created a prototype application for mobile users to chat in three emoji modes: (i) the baseline (i.e., no recommendation), (ii) the current-sentence model, and (iii) the conversation model (*Reeboc*). We recruited 17 participants, 11 male and 6 female (ages 20-29). Using a within-subjects design, we instructed participants to chat about the three topics (plan for a dinner get-together, plan for a weekend group activity, and share recent memorable experiences), each for 10 minutes with the three emoji modes. The task order was fixed, while the order of the emoji modes was randomly assigned for each task.

We analyze which emojis the participants selected during their chat. We hypothesize that with better emoji recommendations, participants are likely to choose emojis that are presented on the top left of the emoji keyboard (as opposed to the bottom right). To measure this effect, we use the `emoji selection rank`, the

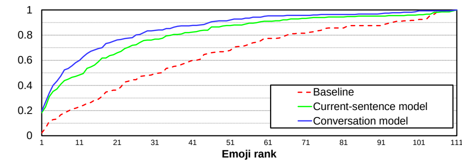


Figure 3: CDF of used emoji "ranks". The X-axis represents the presentation order of the 111 emojis on the keyboard.

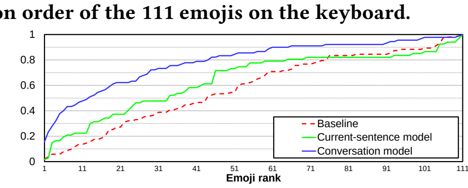


Figure 4: CDF that displays the ranks of selected emojis, specifically for emoji-only sentences.

order in which the chosen emoji is presented in the emoji keyboard. Figure 3 plots the CDF of emoji selection rank comparisons for three modes. Overall, the two recommendation modes showed a considerably better selection rank than the baseline. Comparing between the recommendation modes, the median rank of the CDF is improved from 13 to 7 in the conversation mode over the current-sentence mode.

Our results suggest that emoji-only sentences were used frequently by participants (36.2% of all sentences with emojis). Figure 4 plots the CDF of emoji selection rank comparisons for only emoji-only sentences. Compared to Figure 3, the emoji selection rank difference between the conversation mode and the current-sentence mode is larger.

To summarize, most participants were able to effectively select emoji in the emoji recommendation mode over the baseline emoji layout. The conversation mode generally showed better emoji recommendation than the current-sentence mode, and the emoji selection rank improved by 47%. Especially for the emoji-only messages, the conversation mode showed a much better recommendation, with the emoji selection rank improving by 70%.

With emoji recommendations, participants were able to select emojis faster than when using the baseline. The emoji selection latency in the conversation mode was 2.73 seconds with 1.83 seconds of standard deviation, while the baseline was 4.6 seconds with 4.4 seconds of standard deviation and the current-sentence mode was 3.26 seconds with 2.64 seconds of standard deviation. This suggests that the conversation model improves the latency compared to the baseline by 38% and 14% compared to the current-sentence mode.

4 ACKNOWLEDGMENTS

This research was supported in part by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017M3C4A7083534).

REFERENCES

[1] F. Barbieri, F. Ronzano, and H. Saggion. What does this emoji mean? a vector space skip-gram model for twitter emojis. In *Language Resources and Evaluation conference, LREC*, Portoroz, Slovenia, May 2016.

[2] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.