Technical Disclosure Commons

Defensive Publications Series

March 2020

Sharing photos and associated text messages within messaging apps in the intended order

Victor Cărbune

Thomas Deselaers

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation
Cărbune, Victor and Deselaers, Thomas, "Sharing photos and associated text messages within messaging apps in the intended order", Technical Disclosure Commons, (March 06, 2020)
https://www.tdcommons.org/dpubs_series/2993

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

Sharing photos and associated text messages within messaging apps in the intended order

ABSTRACT

The order in which a photo and messages are delivered and displayed in a chat conversation is affected by processing and/or network delays as well as specific implementations of messaging applications. Receiving a photo out-of-sequence from text messages related to the photo can create confusion and/or misunderstanding and can ruin the intended effect of the sender's communication. This disclosure describes the use of machine learning techniques to identify text messages that are associated with a shared photo and bundling those together as a joint set. Descriptive text for the photo is shown as a placeholder until the photo is fully transferred and displayed. Alternatively, or in addition, messages related to the photo that are sent after the photo is shared are shown to the recipient only after the recipient views the photo.

KEYWORDS

- Chat conversation
- Message order
- Message sequence
- Photo sharing
- Photo caption

BACKGROUND

Users often share photos with others via messaging apps. The user experience (UX) of photo sharing typically includes a user interface that provides users the ability to attach a text caption to the photo being shared. However, a single text caption for a photo is often insufficient to capture all of the text content the user wants to communicate in relation to the photo being shared. In many cases, users send multiple messages related to the photo immediately after

sending the photo. In some cases, users might send a number of text messages related to a photo to build a narrative in anticipation of the photo that is shared subsequently.

It is therefore important that photos and messages related to them be delivered to the recipients in the same order in which the sender sent them. However, the order in which the photo and the messages are delivered and displayed can be affected by various aspects pertaining to network connectivity, such as bandwidth and traffic restrictions, as well as processing delays. Photos are much larger in size compared to text messages and may take a substantially longer time to be transmitted over low-bandwidth networks. Similarly, some networks, such as airplane wireless networks, may block all non-text traffic, thus preventing reception of the photo while letting text messages through. Messaging apps themselves may perform additional processing to send and/or receive photos, thus introducing a delay in delivering and displaying photos compared to text.

Due to the various issues mentioned above, a recipient may receive a photo at a time later than the text messages related to the photo that were sent after the sender shared the photo. Receiving the photo out-of-sequence in relation to text messages related to the photo can create confusion and/or misunderstanding and can ruin the intended effect of the sender's communication. Displaying a low resolution version of the photo while waiting for the full resolution version to transfer or displaying portions of the photo as received does not satisfactorily address this problem.

DESCRIPTION

This disclosure describes the application of machine learning to identify text messages associated with a shared photo and bundling them together as a joint set. The techniques are

implemented with user permission to access and analyze the shared photo and text. Users are provided options to control such access, including providing no access.

When the user permits, content of the messaging conversation is analyzed by the use of one or more of two complementary trained machine learning models:

1. **Image-to-text model:** The shared photo is analyzed by an image-to-text model to generate a short text description that can be used as a placeholder. For example, a hiking photo may result in producing a descriptive text such as "photo with mountains while hiking the Mt. Whitney trail." Such a model can be implemented as a combination of a traditional convolutional neural network (CNN) with Residual Network (ResNet) and Long Short-Term Memory (LSTM) decoder, with the ResNet layer processing the photo and the LSTM decoder outputting the text description word-by-word like a regular seq2seq model. The model can be trained with existing image captioning databases and/or other suitable labeled data.
2. **Image-text association model:** Text messages that are sent after sharing a photo (e.g., within a short time interval) are processed by a model that outputs a score for each message between 0 and 1. The score indicates the likelihood of that message being associated with the shared photo. If the score is above a threshold value, the message is bundled with the photo to create a sharing dependency. Similar to the image-to-text model, the trained image-text association model can consist of an image processing model, such as ResNet. Additionally, it includes an LSTM encoder to embed the text message in a fixed dimensional space. Outputs from these two network components are joined via feed-forward neural network layers to generate a score that indicates the extent to which the message text is related to the shared photo. With

permission from relevant parties, datasets of photos that were shared along with associated text, such as captions, can be used to train the model.

The two trained machine learning models can be used together or separately to improve the user experience of sharing photos via messaging apps. Descriptive text for the photo generated by the image-to-text model can be sent as a text message that is shown as a placeholder until the photo is fully transferred and displayed, thus allowing all subsequent text messages related to the photo to be displayed immediately even if these arrive prior to the photo. Alternatively, or in addition, messages deemed as related to the shared photo according to the image-text association model can be subjected to a sequencing restriction that requires the photo to be transferred and displayed first. Such a restriction ensures that messages related to a photo that are sent after the photo is shared are shown to the recipient only after the recipient has viewed the photo.

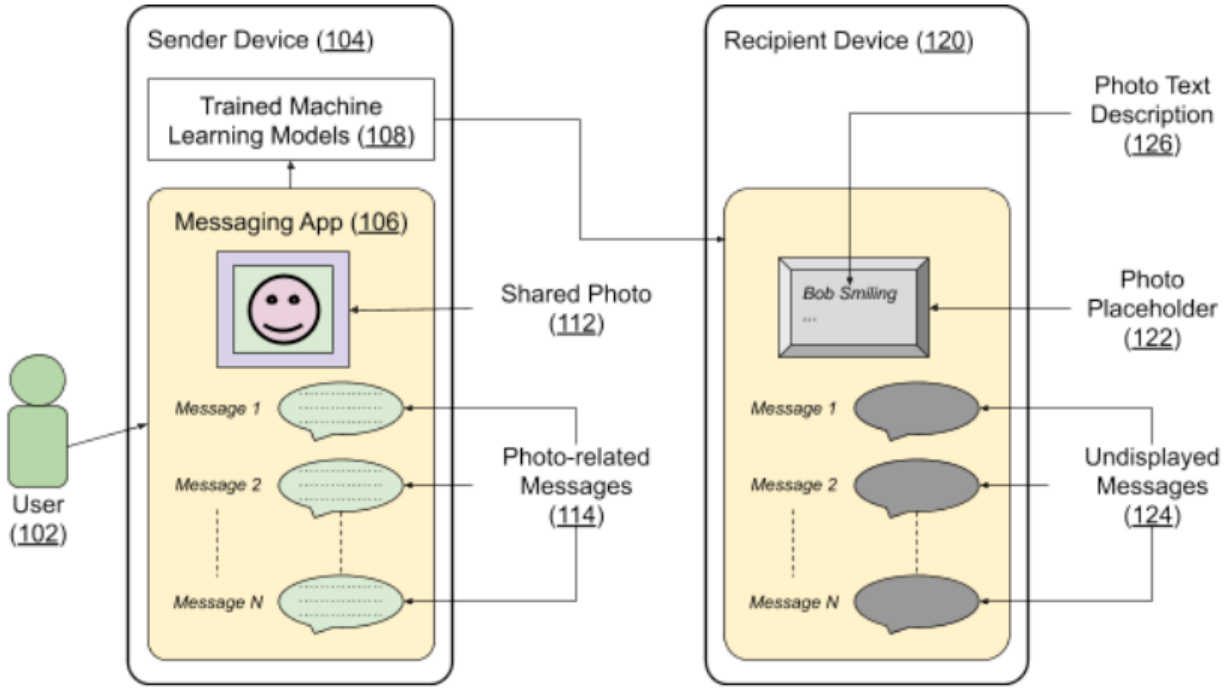

Fig. 1: Delivering a shared photo and related text messages in the intended order

Fig. 1 shows an operational implementation of the techniques described above. A user Bob (102) uses a messaging app (106) on a sender device (104) to share a photo (112) with another user Alice, followed by a set of text messages (114) that are related to the photo. When the user Bob permits, the shared photo and messages are analyzed by trained machine learning models (108) to generate a text description of the photo. The outputs of the models further identifies the text messages as being related to the photo.

The photo, the text messages, and the outputs of the machine learning models are relayed to the messaging app on recipient Alice's device (120). While the photo is being transferred, a placeholder (122) that includes the text description for the photo (126) generated by the machine learning models is displayed. Further, the text messages related to the photo are not shown (124) even if these are received before the photo has finished transferring over the network. Once the photo is received, the placeholder text is replaced with the photo and the text messages related to the photo are revealed, thus ensuring that the receiving user Alice is shown the photo prior to the text messages related to the photo.

The trained machine learning models can operate locally on the sender device as shown in Fig. 1. Alternatively, or in addition, with user permission, the models can be implemented on the receiver device, on a server or any other suitable device. The threshold value to determine whether the association score of a text message is high enough to identify it as being related to the shared photo can be set by application developers, generated dynamically, or specified by the users.

The described techniques can be implemented within any messaging platforms or apps that include the ability to share photos along with text messages. Implementation of the

techniques can enhance the user experience of sharing photos and improve the quality, coherence, and flow of communication.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a user's messages and shared photos, a user's preferences, etc.), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes the use of machine learning techniques to identify text messages that are associated with a shared photo and bundling those together as a joint set. Descriptive text for the photo is shown as a placeholder until the photo is fully transferred and displayed. Alternatively, or in addition, messages related to the photo that are sent after the photo is shared are shown to the recipient only after the recipient views the photo. The techniques can enhance the user experience of sharing photos within messaging apps and improve the quality, coherence, and flow of communication.