Technical Disclosure Commons

Defensive Publications Series

October 16, 2018

Systems and Methods for User Feedback

Lixia Liu

Cassandra Xia

SYSTEMS AND METHODS FOR USER FEEDBACK

Introduction:

The present disclosure provides systems and methods for identifying user input to a computing system, such as to identify when a user is dissatisfied with a service provided by a computing system (e.g., dissatisfied with query results from a search service, video quality of a video-conferencing service, etc.). Typically, there does not exist a natural way for the user to provide input expressing dissatisfaction when a computing system (e.g., a service and/or associated computing device) does not respond as expected by the user. This can make the system opaque and frustrating for the user resulting in a poor user experience. The present disclosure enables the capture of user dissatisfaction in an unobtrusive way. For example, the user can perform a "clap" gesture to signal user dissatisfaction, and the system can detect when the "clap" gesture is performed in order to capture the user dissatisfaction in an unobtrusive way. Furthermore, the system can collect valuable training data associated with the "clap" gesture that can be used to improve the system, such as by training a machine-learned model associated with a service.

Summary:

According to aspects of the present disclosure, when a user is dissatisfied with the behavior of a computing system, the user can provide a predetermined input (e.g., clap, slap, touch, tap, rub, swipe, spoken hotword, etc.) to signal the user's dissatisfaction to the computing system. For example, if the user is dissatisfied with a speech to text rendering of what the user said, the quality of a video call, or the relevance of query results, then the user can provide the predetermined input. In this way, the user can signal the user's dissatisfaction, vent frustration, and provide input to the system that the user viewed the system's performance as less than

satisfactory. Additionally, when the system detects the predetermined input, the system can capture the user's dissatisfaction and attempt to immediately address the cause of the dissatisfaction (e.g., use an alternative speech to text rendering approach, drop the sending and receiving resolution, resync audio/video channels, reduce echo/noise, allocate additional/different resources, output a different set of query results, etc.).

In some implementations, the user can provide additional information relating to the user's dissatisfaction subsequent to the predetermined input. For example, the user can say what the user found dissatisfying about the speech to text rendering, video call, or search results. In this way, the system can collect valuable in-the-moment data that can be used to train and improve the system in the future.

Detailed Description:

Figure 1 depicts a block diagram of an example computing system 100 for identifying when a user is dissatisfied with a service according to example embodiments of the present disclosure. The system 100 includes a computing system 102, a server computing system 130, and a training computing system 150 that are communicatively coupled over a network 180.

The computing system 102 can be any type of computing device, such as, for example, a personal computing device (e.g., laptop or desktop), a mobile computing device (e.g., smartphone or tablet), a gaming console or controller, a wearable computing device, an embedded computing device, or any other type of computing device. The computing system 102 includes one or more processors 112 and a memory 114. The one or more processors 112 can be any suitable processing device (e.g., a processor core, a microprocessor, an ASIC, a FPGA, a controller, a microcontroller, etc.) and can be one processor or a plurality of processors that are operatively connected. The memory 114 can include one or more non-transitory computerreadable storage mediums, such as RAM, ROM, EEPROM, EPROM, flash memory devices, magnetic disks, etc., and combinations thereof. The memory 114 can store data 116 and instructions 118 which are executed by the processor 112 to cause the computing system 102 to perform operations.

In some implementations, the computing system 102 can store or include one or more machine-learned models 120. For example, the machine-learned models 120 can be or can otherwise include models such as neural networks (e.g., deep neural networks) or other types of machine-learned models, including non-linear models and/or linear models. Neural networks can include feed-forward neural networks, recurrent neural networks (e.g., long short-term memory recurrent neural networks), convolutional neural networks or other forms of neural networks.

In some implementations, the one or more machine-learned models 120 can be received from the server computing system 130 over network 180, stored in the computing device memory 114, and then used or otherwise implemented by the one or more processors 112. In some implementations, the computing system 102 can implement multiple parallel instances of a single machine-learned model 120.

Additionally or alternatively, one or more machine-learned models 140 can be included in or otherwise stored and implemented by the server computing system 130 that communicates with the user computing device 102 according to a client-server relationship. For example, the machine-learned models 140 can be implemented by the server computing system 140 as a portion of a web service. Thus, one or more models 120 can be stored and implemented at the computing device 102 and/or one or more models 140 can be stored and implemented at the server computing system 130.

The computing device 102 can also include one or more user input component 122 that receives user input. For example, the user input component 122 can be a touch-sensitive component (e.g., a touch-sensitive display screen or a touch pad) that is sensitive to the touch of a user input object (e.g., a finger or a stylus). The touch-sensitive component can serve to implement a virtual keyboard. Other example user input components include a microphone, a traditional keyboard, camera device, or other means by which a user can provide user input.

The server computing system 130 includes one or more processors 132 and a memory 134. The one or more processors 132 can be analogous to the one or more processors 112 and the memory 134 can be analogous to the memory 114. The memory 134 can store data 136 and instructions 138 which are executed by the processor 132 to cause the server computing system 130 to perform operations.

In some implementations, the server computing system 130 includes or is otherwise implemented by one or more server computing devices. In instances in which the server computing system 130 includes plural server computing devices, such server computing devices can operate according to sequential computing architectures, parallel computing architectures, or some combination thereof.

As described above, the server computing system 130 can store or otherwise include one or more machine-learned models 140. Example machine-learned models include neural networks or other multi-layer non-linear models. Example neural networks include feed forward neural networks, deep neural networks, recurrent neural networks, and convolutional neural networks.

The computing device 102 and/or the server computing system 130 can train the models 120 and/or 140 via interaction with the training computing system 150 that is communicatively coupled over the network 180. The training computing system 150 can be separate from the server computing system 130 or can be a portion of the server computing system 130.

The training computing system 150 includes one or more processors 152 and a memory 154. The one or more processors 152 can be analogous to the one or more processors 112, and the memory 154 can be analogous to the memory 114. The memory 154 can store data 156 and instructions 158 which are executed by the processor 152 to cause the training computing system 150 to perform operations. In some implementations, the training computing system 150 includes or is otherwise implemented by one or more server computing devices.

The training computing system 150 can include a model trainer 160 that trains the machine-learned models 120 and/or 140 stored at the user computing device 102 and/or the server computing system 130 using various training or learning techniques, such as, for example, backwards propagation of errors. In some implementations, performing backwards propagation of errors can include performing truncated backpropagation through time. The model trainer 160 can perform a number of generalization techniques (e.g., weight decays, dropouts, etc.) to improve the generalization capability of the models being trained.

The network 180 can be any type of communications network, such as a local area network (e.g., intranet), wide area network (e.g., Internet), or some combination thereof and can include any number of wired or wireless links. In general, communication over the network 180 can be carried via any type of wired and/or wireless connection, using a wide variety of communication protocols (e.g., TCP/IP, HTTP, SMTP, FTP), encodings or formats (e.g., HTML, XML), and/or protection schemes (e.g., VPN, secure HTTP, SSL).

Figure 1 illustrates one example computing system that can be used to implement the present disclosure. Other computing systems can be used as well. For example, in some implementations, the user computing device 102 can include the model trainer 160 and the training dataset 162. In such implementations, the models 120 can be both trained and used locally at the user computing device 102. In some of such implementations, the user computing device 102 can implement the model trainer 160 to personalize the models 120 based on user-specific data.

According to aspects of the present disclosure, the computing system 102 can provide a service such as processing queries from a user, videoconferencing, etc. For example, the computing system 102 can receive a query from a user (e.g., via input component 122) and output one or more query results (e.g., via output component 124). In some implementations, the input component 122 can include a microphone device and the system 102 can use the microphone device to detect an audible query spoken by a user. In some implementations, the input component 122 can include a camera device and the system 102 can use the camera device to detect a visual query by the user, and/or include one or more input devices (e.g., a keyboard, mouse, etc.) that can be used to receive a query input by the user. In some implementations, the output component 124 can include one or more output devices (e.g., speaker, display, etc.) to output the query results. Upon receiving the query from the user, the system 102 can process the query to determine one or more query results, and output the query results for the user. As another example, the computing system 102 can receive a request to host a videoconference between the user and another user. In some implementations, the input component 122 can include a camera device and the output component 124 can include display and speaker devices. The computing system 102 can use the input component 122 and output component 124 to host the videoconference between the user and the other user (e.g., between the computing system 102 and a computing system associated with the other user, via network 180).

According to aspects of the present disclosure, when the system 102 is providing a service, such as processing a query, hosting a videoconference, performing speech-to-text conversion, etc., the system 102 can monitor for a trigger event that signals the user's dissatisfaction with the system 102. For example, the input component 122 can include a microphone, and the system 102 can use the microphone to monitor for an audible trigger as a result of an audible input performed by the user, such as a clap, slap, or a hotword spoken by the user. As another example, the audible trigger can include the user tapping or rubbing a microphone device of the computing system. As another example, the input component 122 can include a tactile sensor, and the system 102 can use the tactile sensor to monitor for a tactile trigger as a result of a tactile gesture performed by the user (e.g., touch, tap, swipe, etc.). As another example, the system 102 can use a camera device to monitor for a visual trigger as a result of a visual gesture performed by the user (e.g., a hand gesture, facial gesture, covering the camera, etc.). The system 102 can capture the trigger event as a signal that the user is dissatisfied with the system.

According to aspects of the present disclosure, the system 102 can collect in-the-moment data associated with the user's dissatisfaction with the system. For example, if the system 102 detects a trigger event associated with a query, then system 102 can collect data indicative of the query associated with the trigger event, the query results associated with the trigger event (e.g., the query results for the query associated with the trigger event), and/or the trigger event (e.g., a type of the trigger event, a timing of the trigger event, etc.). In some implementations, the system 102 can collect data indicative of a session that includes the query associated with the trigger event. The session can include one or more queries from the user prior to the trigger event, and one or more query results associated with the one or more queries. As another example, if the system 102 detects a trigger event associated with a speech-to-text service, then system 102 can collect data indicative of the input speech, output text, and/or the trigger event. In some implementations, the user can provide additional information about the user's dissatisfaction subsequent to the trigger event, and the system 102 can capture the feedback. For example, the user can say or input what the user found dissatisfying about a speech to text rendering, video call, search results, etc., and the system 102 can associate the user's feedback with the trigger event. In this way, the system 102 can collect highly real-time data where users have directly annotated the specific problematic interaction with the system. This feedback can be more specific and useful for improving the system than general user experience surveys. The system 102 can store the data associated with the user's dissatisfaction and/or the additional information about the user's dissatisfaction as part of the training data 162.

According to aspects of the present disclosure, the computing system 102 can attempt to immediately improve the user experience based on the trigger event. As an example, if the system 102 detects the trigger event when the system 102 is playing music in response to the user's query, then the system 102 can play a different track (e.g., next track) in response to the trigger event. As another example, if the system 102 detects the trigger event when the system 102 is outputting a first ranked element based on query results including a ranked list of elements, then the system 102 can output a second ranked element from the ranked list of elements in response to the trigger event. As another example, if the system 102 detects the trigger event when the system 102 is outputting first query results based on an interpretation of the user's query, then the system 102 can output second query results based on a second interpretation of the query in response to the trigger event. In some implementations, the system 102 can attempt to immediately improve the user experience based on the user's feedback subsequent to the trigger event. As an example, if the system 102 detects the trigger event and collects the user's feedback that indicates that the user is dissatisfied with the syncing of audio and video in a video call, then the system 102 can resync the audio and video in response to the user's feedback. As another example, if the system 102 detects the trigger event and collects the user's feedback that indicates that the user is dissatisfied with the latency of a video call, then the system 102 can drop the sending and receiving resolution of the video call.

According to aspects of the present disclosure, the computing system 102 can attempt to improve the user experience through training based on the trigger event. For example, the system can collect real-time, labeled training data 162 for improving the system based on the trigger event and/or the user's feedback subsequent to the trigger event. In particular, the model trainer 160 can train the machine-learned models 120 and/or 140 based on a set of training data 162. The training data 162 can include, for example, the data associated with the user's dissatisfaction and/or the additional information about the user's dissatisfaction.

As an example, the model trainer 160 can train machine-learned models 120 and/or 140 associated with determining relevant query results based on training data 162 that includes data indicative a user's dissatisfaction with query results. As another example, the model trainer 160 can train machine-learned models 120 and/or 140 associated with speech to text based on training data 162 that includes data indicative of a user's dissatisfaction with a speech to text rendering of what the user said. As another example, the model trainer 160 can train machine-learned models 120 and/or 140 associated with hosting and/or setting-up a videoconference based on data indicative of a user's dissatisfaction with a videoconference.

Drawings:

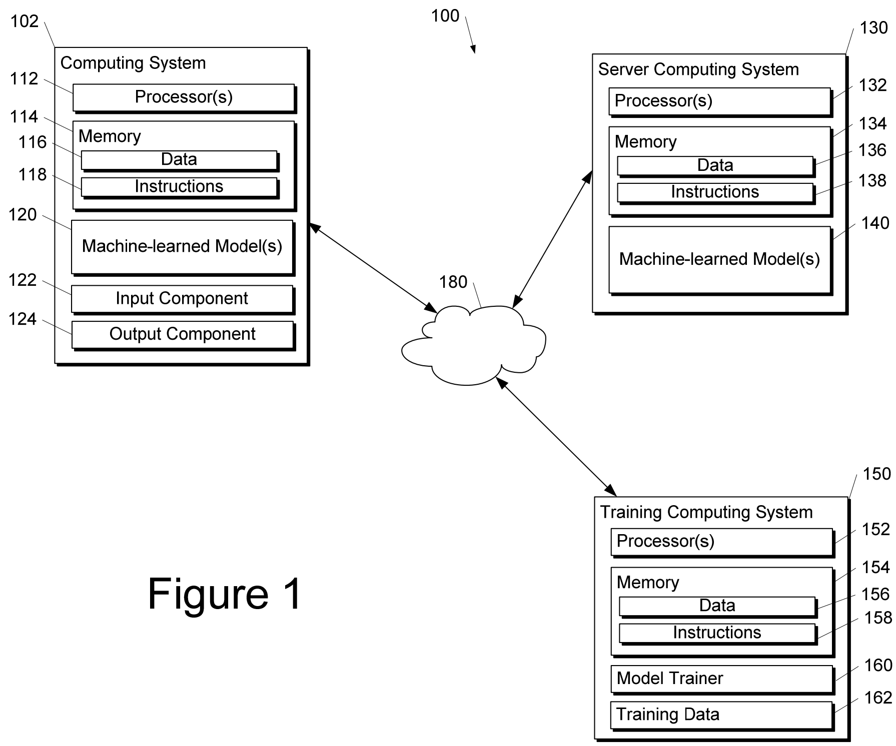

Figure 1

Abstract:

The present disclosure describes systems and methods for identifying user input to a computing system, such as to identify when a user is dissatisfied with a service. According to aspects of the present disclosure, a computing system that is providing the service can monitor for a trigger event that signals a user's dissatisfaction with the system. The system can collect in-the-moment data associated with the user's dissatisfaction with the system, and attempt to immediately improve the user experience based on the trigger event and/or attempt to improve the user experience through training based on data associated with the trigger event. Keywords associated with the present disclosure include: computing systems (e.g., smartphone, smartwatch, mobile phone, automated assistant, home assistant, personal assistant); user; user experience; user feedback; user dissatisfaction; trigger event; hotword; query; query results; gesture recognition.