February 14, 2019

Providing Responses To Queries Received In A Group Context

Anonymous

Providing Responses To Queries Received In A Group Context

ABSTRACT

Techniques to provide responses to user queries are described herein. Sensor data from a client device is obtained with user permission to detect the presence of users near the device. Such users can include verified users and unverified users. User profiles for the users are accessed to retrieve information regarding the group, such as a total number of users, age/gender information, mood information, etc. The query context is determined, e.g., location and time of the query, weather at the location, etc. The user query is analyzed using natural language processing techniques. A customized response is generated based on the retrieved user profiles and the query context and is provided via the client device. The described techniques can be implemented as a standalone application, or as part of other applications such as a social networking application or a messaging application.

KEYWORDS

- Virtual assistant
- Smart assistant
- Smart speaker
- Query context
- Group presence
- User profile
- Conversational agent
- Bot

BACKGROUND

Users utilize devices such as smartphones, tablets, smart speakers, home video-conferencing devices, smart appliances, etc. to perform various tasks. Many devices include features that enable users to ask a query and have the device provide a response or perform an action in response to the query. For example, a user can ask a smart speaker in the kitchen for a recipe, ask a smartphone to place a call, etc. Providing appropriate responses to user queries is an important function of such client devices. For example, client devices can incorporate software programs, e.g., a virtual assistant, or a messaging application, that answers queries spoken by the user or input via an input device such as a device keyboard. Queries can include text, speech, images, etc.

Responses to user queries are based on a combination of the user query and contextual factors such as location. Information from online sources (e.g., weather, traffic, news, stock market information, calendars, etc). Responses can include providing information that answers a user question; performing a task for the user, e.g., booking a restaurant, purchasing concert tickets; managing the user's calendar, e.g., updating an event time if the user is running late; etc.

In certain situations, a user may provide a query in the presence of other users. The user that provides the query may be a known user (e.g., device owner), or an unknown user. In these situations, where a group of two or more users is present, responses determined without taking into account presence of the other users may not be optimal or may be unsuitable for the group.

DESCRIPTION

This disclosure describes techniques to provide customized responses to user queries received in the presence of multiple users, based on context. For example, user queries can include voice queries, received via a microphone of a client device such as a smartphone, tablet, smart speaker, video-conferencing device, home appliance, or other device.

Fig. 1 illustrates an example process to generate a customized response to a user query. The process can be implemented, e.g., by the client device, by a server device that is accessible from the client device over a network, or a combination. For example, the process can be implemented as part of a standalone assistant application, or as part of other applications such as a social networking application, a messaging application, etc. Responses to user queries can be provided, e.g., by a conversational agent or chat bot.

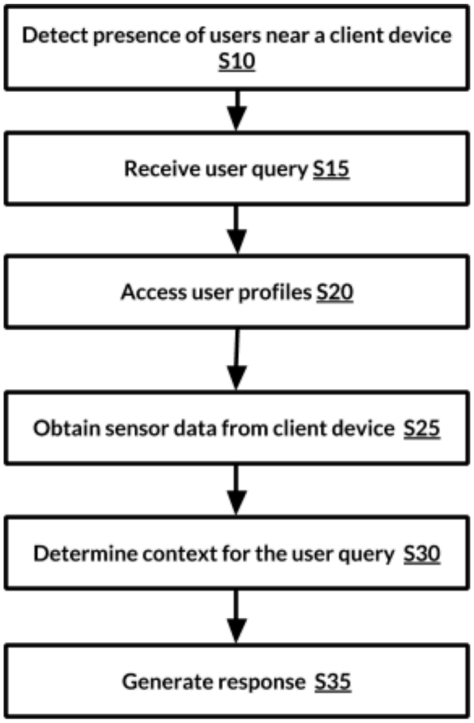

Fig. 1

Detect users present near the client device (S10)

Presence of users near the client device is detected using various sensors. For example, such sensors can include a device camera, microphone, infrared sensor, temperature sensor, motion sensor, etc. Data from such devices is analyzed using techniques such as computer vision

techniques (e.g., to process data from the camera) and various signal processing techniques to process other data.

The users present near the client device can include verified users, e.g., users whose identity is known, and unverified users. In different situations, the users present near the device can include all verified users, all unverified users, or a mix of verified and unverified users. Presence of the users is detected using data from various sensors of the client device.

Receive user query (S15)

A user query is received at the client device. The user query can be received from any user near the client device, e.g., a verified user or an unverified user. For example, a user can provide the query as spoken phrase, typed input, image, video, etc. Multimodal query input is supported. Prior queries received on the client device can be stored and utilized, if permitted by the user. A conversation format is used to indicate received queries and responses provided.

For example, the spoken phrase can include an activation phrase such as “Assistant” followed by the actual query. The client device detect the query using the microphone or other sensor. Text queries, e.g., typed using a keyboard, can be processed directly, while audio or video queries can be processed using speech recognition techniques to convert the query into text.

For example, the user query can be processed, e.g., using natural language techniques, to determine that the query is a request for a recommendation. The query can be analyzed to detect user intent (e.g., “Play music,” “Open app,” etc.), domain (e.g., “songs,” “videos,” etc.), individual items in the query (e.g., “Beethoven’s Fifth Symphony,” “Coffee shop X.” etc.) etc. Knowledge representations of entities and their relations can be utilized in interpreting the query.

Entities can include users, concepts, businesses, object names, etc. and each entity can have associated attributes.

Access user profiles (S20)

User profiles of the users detected near the client device can be accessed. For example, user profiles for verified users can be accessed based on the determined identify of the verified users. Further, when the users detected near the client device include unverified users, a candidate profile of the unverified users is accessed. User profiles can include user information, e.g. demographic information, and context information associated with the user. User profiles and candidate profiles can include information regarding corresponding users, such as biographic data; demographic data (such as age, gender, etc.); user interests and preferences; user location; educational and work information; memberships and connection with other users; etc.

Entity resolution techniques can be used to identify candidate profiles. For example, if an unverified user is detected, entity resolution techniques can be applied to match to an available candidate profile for the unverified user based on available sensor data. Based on the entity resolution, relationships between a verified user and the unverified user can be determined by analyzing the respective profiles. In some instances, it may not be possible to determine the relationship using such techniques. In such cases, candidate profiles are used without an explicit association with a known entity.

The user profiles and candidate profiles can be stored on a server that maintains user profiles. Candidate profiles can be stored in association with the user profile of a verified user and made accessible when the user profile of the verified user is accessed. User profiles can be associated with privacy settings, and users associated with individual profiles can limit access to the profiles. User profiles are accessed based on the associated privacy settings.

Candidate profiles of unverified users can be constructed based on sensor data associated with the client device. Such data can include data measured in the background, e.g., when the client device is not in active use for a user query, prior to receiving the query. The candidate profiles can include information regarding prior interactions between a verified user and the client device when the particular unverified user was present near the client device. The candidate profiles can also include prior interactions between the unverified user and the client device. Prior interactions can include, e.g., user queries received by the client device, prior to the current user query. Candidate profiles can be converted to user profiles, e.g., when an unverified user is verified subsequent to creation of the candidate profile.

Obtain sensor data (S25)

Sensor data is obtained from sensors of the client device. For example, such sensors can include a device camera, microphone, infrared sensor, temperature sensor, motion sensor, etc. The client device may be configured to periodically or continuously obtain data from such sensors, e.g., prior to and during the time that the query is provided. Such data capture is based on permission provided by the users.

The obtained sensor data is processed to detect users that are near the client device, to identify the user that provided the query, to predict the users that receive the response (e.g., listen to a spoken response provided by the client device), etc. Detection of users based on the sensor data can include determination of the number of users near the client device, various attributes such as gender, age, etc.

Determine context for the user query (S30)

The context of the user query is determined. Determination of the context is based on the data obtained from device sensors. The context can include various factors such as a total number of users detected near the client device, an age or gender profile of the group of users detected near the client device, a detected mood of the users, etc.

The determined context can also include the geographic location for the query, e.g., the location of the client device at the time the query is received, and information associated with the geographic location, e.g., weather at the location. The context can also include factors determined based on user profiles of verified users and candidate profiles of unverified users that are present near the client device at the time the user query is received. For example, the context can include a characterization of the group, e.g., "family," "roommates," "co-workers," "college students," etc.

Generate response (S35)

A customized response to the user query is generated. The response is based on profiles of the users near the client device and the determined context. The response is personalized based on these factors. The response can be customized based on relationships between the various users that are present. Responses to similar user queries are thus different, based on the context. Responses can be generated by an agent, e.g., provided as part of an application that receives the query. Responses can also be generated by interacting with different agents or bots that are associated with entities that match the user query. For example, requests for ticket reservations can be handled by a third-party reservation system, requests for music playback can be handled by a music service associated with the client device (e.g., a service to which the user of the client device has a subscription), etc.

The customized response can be generated, e.g., by a server. The customized response is presented to the user. For example, the response can be provided as a spoken response using a speaker of the client device. Presentation of the response can also be personalized based on the user profiles or the context. The response can be presented using natural language generation techniques. For example, the response can be provided as a spoken response using text-to-speech techniques.

The techniques to provide responses to user queries can be implemented as part of a virtual assistant application that is provided by a social network or a messaging application. For example, the query responses can be provided by a bot or assistive agent provided by the social network or messaging application. The bot can assist the user to interact with the social network or messaging application, e.g., to create social network posts, to write comments, to send messages, etc.

The process described with reference to Fig. 1 can be implemented using various techniques, including machine learning. For example, the user query may be parsed by use of a deep learning architecture that includes multiple long-short term memory (LSTM) networks. The query parser can be based on a recurrent neural network grammar (RNNG) model. Entity resolution can be performed using techniques such as support vector machines (SVM), regression models, deep convolutional neural networks (DCNN), etc. Knowledge representation using feature vectors or embeddings can be used.

Example of use

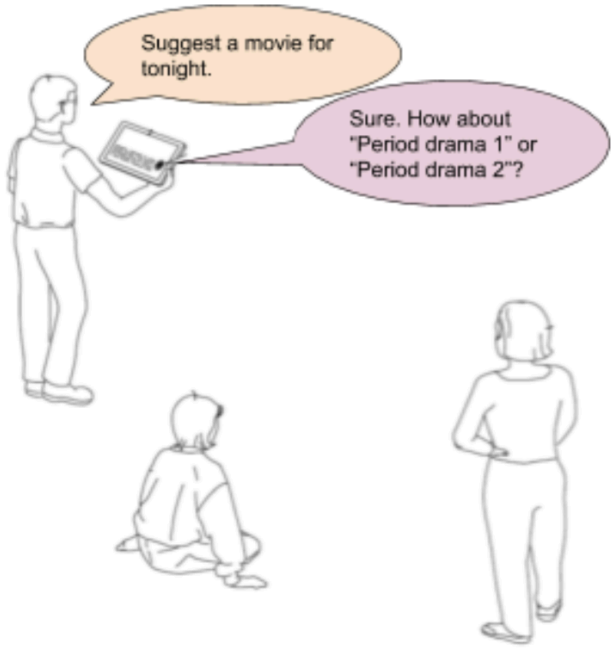

Fig. 2A

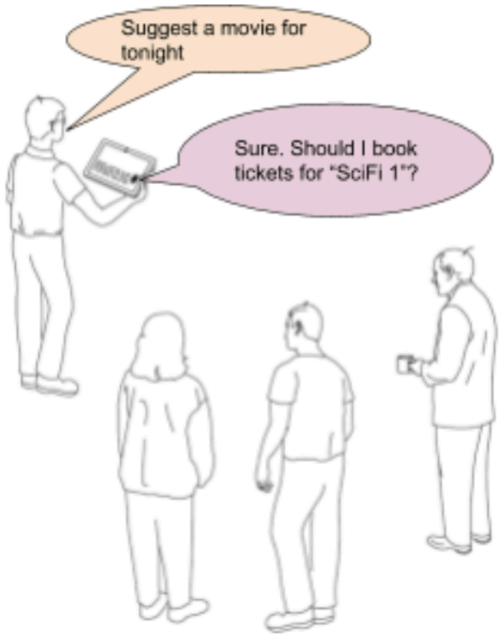

Fig. 2B

Fig. 2 illustrates a user query and corresponding response generated using the techniques described herein. In Fig. 2A, it is determined that the user query ("suggest a movie for tonight") is received while the user is at home. Other users that are detected include the user's family members, determined based on user profiles or candidate profiles. Accordingly, the response includes suggestions for period dramas. For example, period dramas are suggested by taking into account the prior viewing history of the group, available information regarding movies previously watched or rated by individual users near the device, subscription and catalog information for an online movie streaming service, a current time and location, etc.

In Fig. 2B, it is determined that the user query (identical the query in Fig. 2A) is received while the user is at a workplace. It is detected that three other individuals, e.g., co-workers are present. Based on context and user profiles, a suitable response "Should I book tickets for SciFi 1?" is provided. For example, the response is provided based on determination that the query is

received during a workday, that the user's calendar indicates a team outing in the evening, and based on user profiles of the co-workers. As described above with reference to Fig. 1, it is not necessary that all users present near the device be verified users, in order to provide a suitable query response.

CONCLUSION

Techniques to provide responses to user queries are described herein. Sensor data from a client device is obtained with user permission to detect the presence of users near the device. Such users can include verified users and unverified users. User profiles for the users are accessed to retrieve information regarding the group, such as a total number of users, age/gender information, mood information, etc. The query context is determined, e.g., location and time of the query, weather at the location, etc. The user query is analyzed using natural language processing techniques. A customized response is generated based on the retrieved user profiles and the query context and is provided via the client device. The described techniques can be implemented as a standalone application, or as part of other applications such as a social networking application or a messaging application.