Customizing Advertisement Delivery In Voice Activated Computing Systems Based on Status Information Of A Service Provider and Users

John D. Lanza
Foley & Lardner LLP

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

Recommended Citation
Lanza, John D., "Customizing Advertisement Delivery In Voice Activated Computing Systems Based on Status Information Of A Service Provider and Users", Technical Disclosure Commons, (July 03, 2017)
http://www.tdcommons.org/dpubs_series/577

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.
CUSTOMIZING ADVERTISEMENT DELIVERY IN VOICE ACTIVATED COMPUTING SYSTEMS BASED ON STATUS INFORMATION OF A SERVICE PROVIDER AND USERS

Voice activated computing systems provide a user with content or services in response to voice commands spoken by the user. Such systems can capture voice commands from a user, process the voice commands to determine requests and keywords in the voice commands, and provide the user with content or services related to the determined requests and keywords.

As discussed herein, a voice activated computing system processes the voice commands and selects ads to be presented to the user based, in part, on a status of the merchants associated with the ads, and in part on the status of the users. For example, a restaurant may have an open dinner reservation for 7 pm, which the restaurant would like to fill before that time. The restaurant can present a related ad to the system providing details of the open reservation and the expiration time of 7 pm. The system can receive location data from a set of candidate user devices. The system may also receive user preferences for the types of restaurants the user likes or for which it is willing to accept ads. The system can send an ad for the open 7 pm reservation at the restaurant to selected users. The selected users can be located within a certain distance from the restaurant that would reasonably allow the user to get to the restaurant by 7 pm. Further, the selected users may have previously expressed interest in receiving ads for the cuisine that the restaurant offers. The selected users may also include those users that have expressly sent a voice command to the system indicating an interest in going to a restaurant. The system may also determine the frequency with which the user may receive the ad based on the characteristics of the user device, such as battery power, processing power, display settings, etc. For example, if the user device is running low on power, the system may send the ad to the user device only every 10 minutes, instead of every 5 minutes. In this manner, the system can
improve the matching of service providers with users based on the status of the service provider and the status of the users.

![Diagram of voice activated computing system]

Figure 1 shows an example voice activated computing system. The system includes a voice assistant device, a service provider, a content provider, and a data processing system communicating over a network. The voice assistant device can be a device that accepts voice commands, and provides audio or visual output. The voice assistant can include one or more mics and cameras, such that voice commands received by the user are converted into corresponding audio signals. The voice assistant can send the audio signals to the data processing system and the service provider. The voice assistant device also can receive data such as audio signals or video signals from the data processing system or the service provider.
The voice assistant device also can include audio speakers that can convert the audio signals received from the data processing system or the service provider into sound.

The data processing system can process voice commands received from the voice assistant device. The data processing system includes a natural language processor, an audio signal generator, a task predictor, and a content selector. The natural language processor is capable of processing voice commands included in the audio signals received from the voice assistant device. The natural language processor can convert the audio signals into recognized text by comparing the audio signals against a stored, representative set of audio waveforms, and choosing the closest matches. The representative waveforms are generated across a large set of users, and can be augmented with speech samples. After the audio signals are converted into recognized text, the natural language processor can match the text to words that are associated, for example via training across users or through manual specification, with actions that the data processing system can serve. Basically, the natural language processor identifies requests and trigger words in the converted text, based on which the natural language processor can determine the content and actions to be carried out. The task predictor can predict tasks or actions based on the converted text, and in particular by identifying requests and trigger keywords in the converted text. The task predictor also can predict the most likely sequence in which the tasks would be executed. The content selector can select content, such as services to be offered to the user based on the actions identified by the task predictor. In addition, the content selector also can alter the sequence or the order in which the actions related to the services offered to the user are executed. The audio signal generator can generate audio signals based on the services selected by the content selector. The audio signals can be representative of voice responses or voice instructions provided to the user in response to the voice commands.
The service provider can provide one or more service to the user. For example, the service provider can be a taxi or car sharing service provider, dining or reservation service provider, and the like. The service provider can communicate with the voice assistant device independently of the data processing system and provide the user the ability to request a ride, do a dinner reservation, or avail of other services provided by the service provider. The service provider can also include a natural language processor, similar to the one discussed above in relation to the data processing system, to convert user voice commands into text, and identify requests and keywords to determine the services requested by the user. The service provider also can send to the system requests for ads to be presented to users in addition to providing their status to the system. For example, referring to the foregoing example, the service provider can send a request to present an ad to users for an open dinner reservation at the restaurant at 7 pm. In some other examples, the service provider also can provide the user with status information such as whether the service provider is open or closed for business, or the inventory levels of a particular product that the service provider is advertising.

The content provider can provide sponsored content items related to the requests and trigger keywords identified in the voice command. The content items provided by the content provider can be provided to the user in addition to the response to the requests made by the user in the voice command. The content provider also can provide ads referred to by one or more service providers. For example, the content provider can provide the system with the ad for the 7 pm dinner reservation at the restaurant.

The user can speak the voice command “I would like to go out to a restaurant,” to the voice assistant device. The mics at the voice assistant device can convert the voice commands
into corresponding audio signals, which are be transmitted by the voice assistant device to the
data processing system over the network.

At the data processing system, the natural language processor processes the audio signal received form the voice assistant device and identify a request for a “restaurant.” The natural language processor also can identify a trigger keyword “go” or “to go out to,” which can indicate a need for transportation. Even though the user’s voice command does not directly express an intent for transportation, the trigger keyword indicates that transportation may be needed.

The task predictor, based on the requests for “restaurant” and on the trigger keywords, can determine a most likely sequence of actions related to the voice command. For example, the task predictor can determine a task sequence that includes restaurant reservations, booking a ride to the restaurant, and booking a ride from the restaurant.

The content selector can determine the names of the restaurants to be presented to the user. As mentioned above, the content selector can determine which of the many names of restaurants provided by the service providers to be presented to the user based on the status of the restaurant and the status of the users.

The status of the restaurant, with regards to the ad presented by the restaurant, can be presented to the system in a request received from the restaurant (or a restaurant reservation service provider). For example, the request sent by the restaurant can include a time (e.g., 7 pm) or an inventory (e.g., two open tables for two) information. In some implementations, the restaurant may provide the status information only once when the request is sent to the system. In some other implementations, the restaurant can send the status information several times at regular or irregular intervals. The status information of the restaurant may also be determined from historical data related to the restaurant. For example, the content selector can estimate the

Lanza: Customizing Advertisement Delivery In Voice Activated Computing S

Published by Technical Disclosure Commons, 2017
inventory at the restaurant based on prior instances of users accepting reservations at the
restaurant through the system. Based on the status information of a service provider, the content
selector can select which ones of the ads associated with the service provider that the system can
select to present to the users. For example, if the restaurant is currently closed, the content
selector can refrain from sending any ads for reservations to the users. In another example, if the
restaurant has provided an ad for a 7 pm reservation, but also indicates that there is no inventory,
the content selector may refrain from providing the ad to the users.

The current location of the user can be a location of a user device that was used to
transmit the voice command to the system. For example, if the user used a mobile device to send
a voice command to the system, the location of the mobile device can be used as a representative
location of the user. In some implementations, the current location of the user can be a location
of a second user device that is associated with the user device that transmitted the voice
command to the data processing system. For example, multiple devices can be associated with
one another through a unique login, a family account, or other grouping. The system can use the
location of the user device from which the voice command originated or the location of any other
user device associated with the originating user device. In some implementations, the current
location of the user can be determined based on location information communicated from the
user device to the system for the purpose of locating the device (such as geographic coordinates
communicated by the device when the user has granted location-based information to be
communicated). In some implementations, the location can be determined based on network
information associated with the user device. This network information can include, for example,
an IP address of the user device, an IP address of a carrier associated with the user device, or a
network identification name (e.g., an SSID of a nearby wireless network). The location of the
user device may also be manually set by the user. For example, for a home voice assistant device, the user can provide a location to the system to associate with the home voice assistant device. In some implementations, the location of the user device can dynamically update. The location of the user device may update when polled, every minute, every 30 minutes, every hour, or every several hours. The location can be determined using a GPS module within the client device or cellular tower location information.

The user interest can be determined, for example, based on prior requests or voice commands associated with the user. For example, the interest can be based on past search histories provided by the user to the system. The interests can also be based on prior web browsing histories, selection of products or services, selection of sponsored content, or actions associated with the request.

In some instances, the content selector can select a user device that is located closest to the service provider. The selection can be based on a distance from the candidate device to a location associated with the service provider. For example, the first device can be a candidate device that is located closest to the restaurant that has the open reservation. In some implementations, the selection can be made based on a status associated with the candidate devices (or other degree of relationship between a user of the client device and the service provider). For example, the service provider may offer loyalty points when shopping in their retail locations. In this example, the first device can be candidate device whose user has accumulated the most loyalty points at the restaurant. In some implementations, the data processing system can rank the plurality of candidate devices based on the characteristics of the plurality of candidate devices. The first and second device can be identified based on the ranking of the candidate devices. For example, the first and second device can be the candidate devices
with the best network utilization or most remaining battery life. In some implementations, the first device is selected from the candidate devices and the second device is a device that is associated with the same account as the first device. For example, the first device can be the user's smart phone. The second device can be the user's smart assistant device.

In some instances, the content selector can calculate a pooling parameter for the first and the second device. The pooling parameter can indicate a level of overlap between the first device and the second device. For example, the pooling parameter can indicate the first and the second devices are both associated with the same user account. The pooling parameter can also indicate differences or similarities between the first and the second devices. For example, the pooling parameter can indicate which of the first and the second device is closer to the location of the restaurant. The pooling parameter can be determined based on a heuristic technique applied to the characteristic of the first device and the second device. A pooling parameter is calculated for the different pairs of devices in the plurality of candidate devices. The first and second device can be selected from the plurality of candidate devices and the pair with the best (e.g., highest pooling parameter).

The content selector can send the selected ad to the selected user (and selected one of the users devices). If the ad is being sent to the voice assistant device of the user, the content selector can have the audio signal generator to generate an audio signal representative of the ad to be presented to the user. In some instances, the content selector can have the service provider send the audio signal to the user. As mentioned above, the content selector can determine the frequency with which to send the ads to the user device based on the status of the user device.
The user can respond to the ad by accepting the ad, ignoring the ad, rejecting the ad, or requesting additional information. If the system receives a response from the user, the system can process the response appropriately.
Abstract

This document describes a technique for processing voice commands in a voice activated computing system. In particular, the system processes the voice commands and selects ads to be presented to the user based, in part, on a status of the merchants associated with the ads, and in part on the status of the users. The system takes into consideration the status information of the service provider, which status can include whether the service provider is open or closed for business, inventory levels of an advertised product at the service provider, etc. The system also takes into consideration the status information of the user, which status information can include the location of the user relative to a location associated with the service provider, interests of the user, power level of the user device, etc. Based on the status information of the service provider and the status information of the user, the system can select an appropriate ad from the service provider and present the ad to a selected an user from a number of users.