Technical Disclosure Commons

Defensive Publications Series

January 18, 2019

Challenge keyword generation for voice-based authentication

Shuyang Chen

Nicholas Felker

Challenge keyword generation for voice-based authentication

ABSTRACT

This disclosure describes use of challenge keywords for use in voice authentication systems. Verification keywords are generated for a particular user and serve to verify the authenticity of the user. With user permission, keyword generation is performed based on analysis of previous voice-based interactions of the user. A uniqueness score is assigned to words uttered previously by the user in interactions with the voice interface. The uniqueness score for a word is determined based on a rarity of usage of the word in the user's speech and a distinctiveness of user pronunciation of the word.

KEYWORDS

- virtual assistant
- voice UI
- voice fingerprint
- voice authentication
- voice spoofing
- challenge response
- smart speaker
- audio spoofing

BACKGROUND

Voice recognition and response technologies are increasingly integrated into consumer devices. These devices are designed to receive and respond to voice commands respond to voice. Such devices include smartphones, smart speakers, smart appliances, internet-of-things (IoT) devices, etc. Although convenient and intuitive to use, use of voice recognition technologies open up the possibility of malicious actors fraudulently using the voice interface to perform unauthorized actions.

DESCRIPTION

This disclosure describes use of keywords for use with voice based CAPTCHAs and other voice authentication systems. Per techniques of this disclosure, a verification phrase is generated for a particular user that serves to verify the authenticity of the user. With user permission, a history of user transcripts and audio clips from past user interactions with a voice-based interface, e.g., as provided, by a virtual assistant, is utilized in the generation of the verification phrase.

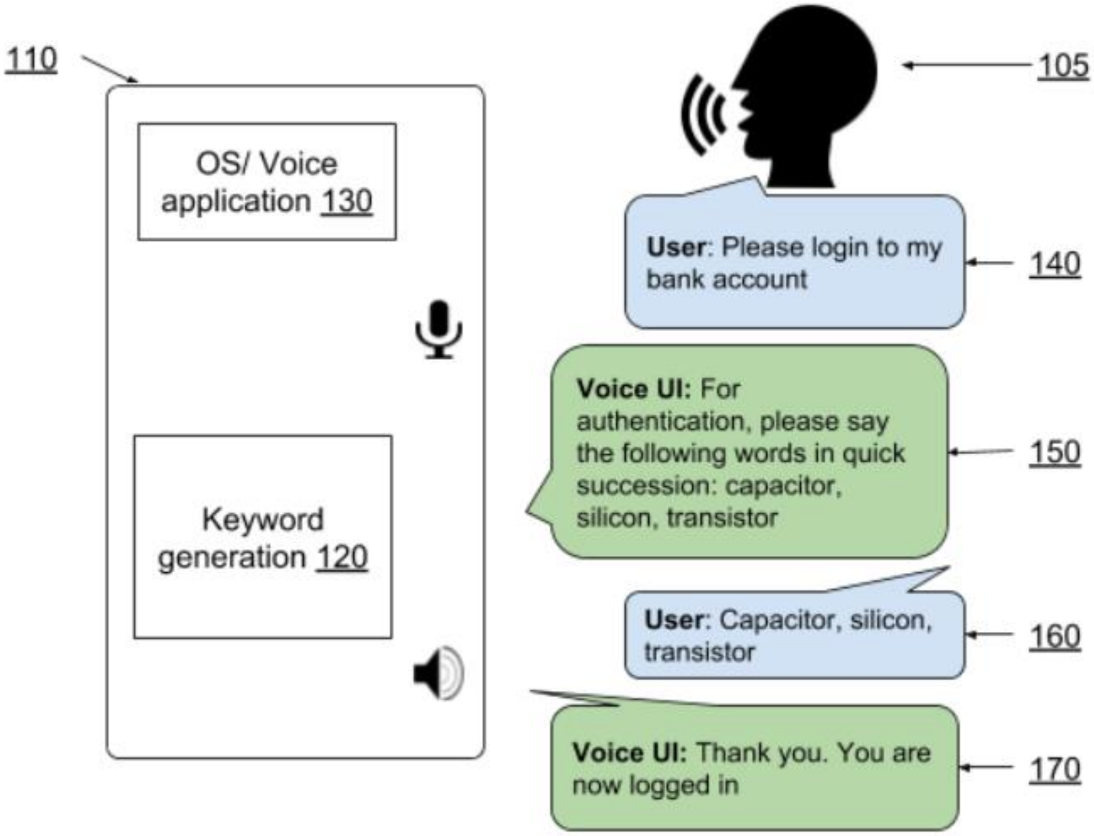
110
OS/ Voice application 130
Keyword generation 120
105
User: Please login to my bank account
140
Voice UI: For authentication, please say the following words in quick succession: capacitor, silicon, transistor
150
User: Capacitor, silicon, transistor
160
Voice UI: Thank you. You are now logged in
170

Fig. 1: Keywords for authentication are tailored to the user to prevent spoofing

Fig. 1 illustrates an example of user authentication via a voice user interface, per techniques of this disclosure. In this illustrative example, a user (105) requests access (140) to their bank account using a voice interface application (130) provided by user device (110).

The voice interface application responds by issuing a challenge (150) wherein the user is requested to authenticate their voice with utterance of certain specified challenge keywords. In this illustrative example, the challenge keywords 'capacitor,' 'transistor,' and silicon' are generated by a keyword generation application (120). The keyword generation application can be implemented on the user device (as shown) and/or on a server.

Upon receipt of the user response (160), the voice interface application tests the response of the user for authenticity by utilizing a voice fingerprint model that is generated from past user interaction. A comparison of the user response with previously recorded speech stored in the voice fingerprint model is used to generate a match score. Determination of successful authentication is made when the match score meets a threshold. A suitable response (170) is provided upon the determination of successful authentication.

With user permission, keyword generation is performed based on analysis of previous voice-based interactions of the user. A uniqueness score is assigned to words uttered previously by the user in interactions with the voice interface. The uniqueness score for a word is determined based on rarity of usage of the word in the user's language and a distinctiveness of the user pronunciation of the word. User accent can lead to distinct pronunciations for certain words compared to the manner in which the words are pronounced by other users. Words that sound notably different due to user accent are assigned a high uniqueness score.

The uniqueness score of words for a particular user can change over time as other users interact with the voice interface. For example, additional users who also use certain uncommon words can cause such words to drop in their uniqueness score, new users with certain accents can lead to a drop in uniqueness scores for words spoken by other users who also speak with those accents. Accordingly, the list of challenge words used for verification of particular user is updated over time depending on a recent vocabulary of the user. The update is reflective of recent accent changes.

Techniques disclosed herein can prevent audio spoofing. The use of an individual’s unique speech pattern renders it difficult for a malicious actor to attack a voice-based user interface by providing a spoofed response to the keyword challenge.

Alternatively or in addition to user-specific keywords, the challenge keywords can include random words that are not specific to the particular user. Such random words are generated without taking into account an individual’s usage patterns. For example, the random words are selected based on an identification of a certain number (for example, three or more) of words that do not appear together frequently in speech transcripts.

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 social network, social actions or activities, profession, a user’s preferences, or a user’s current location), 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 use of challenge keywords for use in voice authentication systems. Verification keywords are generated for a particular user and serve to verify the authenticity of the user. With user permission, keyword generation is performed based on analysis of previous voice-based interactions of the user. A uniqueness score is assigned to words uttered previously by the user in interactions with the voice interface. The uniqueness score for a word is determined based on a rarity of usage of the word in the user’s speech and a distinctiveness of user pronunciation of the word.