Acoustic Fingerprints for Access Management in Ad-Hoc Sensor Networks

Pablo Pérez Zarazaga
07 September 2020
General outline

- Popularity of speech interfaces
- Wireless acoustic sensor networks
- Privacy concerns
- Audio-based device pairing
- Proposed methods
- Results
- Conclusion
Popularity of Speech Interfaces

Alexa, order a pizza for dinner

Do you want your usual order from Alfredo´s?

Yes, please
Popularity of Speech Interfaces

Alexa, order a pizza for dinner

...
Wireless acoustic sensor networks

Image extracted from http://pixabay.com
Wireless acoustic sensor networks

- Many electronic devices around us.
- We have methods to synchronize multiple devices in a network.
- They could collaborate together to provide a distributed voice user interface.
Privacy concerns

- Multiple devices recording and sharing our voice information

- Our voice contains a great amount of personal information.

- How can we know which devices in the network can be trusted to process our voice?
Privacy concerns: Example

Alice

Bob
Privacy concerns: Example
Proximity-based device pairing

- File sharing applications use Bluetooth signals to detect the proximity of other devices.
  - Google’s Nearby Share
  - Apple’s AirDrop
Proximity-based device pairing

- File sharing applications use Bluetooth signals to detect the proximity of other devices.
- Physical proximity is not reliable in a conversational setting.
Proximity-based device pairing

- File sharing applications use Bluetooth signals to detect the proximity of other devices.
- Physical proximity is not reliable in a conversational setting.
- Solution: Adapt the authentication to our perception of privacy
How can we protect our privacy?

- In order to protect the user’s privacy, we need to understand how people perceive their privacy.
How can we protect our privacy?

- People modify the way they talk depending on the privacy of the environment.
- We can use the audio of the environment to recognize devices in proximity.
Acoustic fingerprints

- An acoustic fingerprint compresses a set of features of the audio signal to allow an easy comparison of audio segments.

- A popular application of acoustic fingerprints is music retrieval.
  - For example: Shazam
Acoustic fingerprints

- An acoustic fingerprint compresses a set of features of the audio signal to allow an easy comparison of audio segments.

---

Aalto University
School of Electrical Engineering
Acoustic fingerprints

- An acoustic fingerprint compresses a set of features of the audio signal to allow an easy comparison of audio segments.
Audio-based device authentication

Eve

Reach of Bob's voice

Fingerprint Mismatch

Alice

Fingerprint Match

Bob

Aalto University
School of Electrical Engineering
Audio-based device authentication

- Previously presented methods:
  - Mix of audio fingerprints and fuzzy cryptography.
  - Long non-overlapping windows against de-synchronization.
  - The required audio recording is too long for a conversational application.
Audio-based device authentication

- Previously presented methods:
  - Mix of audio fingerprints and fuzzy cryptography.
  - Long non-overlapping windows against de-synchronization.
  - The required audio recording is too long for a conversational application.

- We need to reduce the length of the audio recordings.
Solution

- Use similar parameters to typical speech processing applications:
  - Shorter windows
  - Overlap
Solution

- Use similar parameters to typical speech processing applications:
  - Shorter windows
  - Overlap

- We propose decorrelation methods to compensate for the degradation in the windowing process.
Decorrelating transformations

- Eigenvalue decomposition
- Wiener filtering
- 2D - DCT
Decorrelating transformations

- Eigenvalue decomposition

- Wiener filtering

- 2D - DCT
Quantization methods

- **Entropy-based:**

- **Mutual-information-based:**
Experiments

- Simulated conversation between two speakers.

- Conversation simulated in multiple realistic scenarios.
## Fingerprint parameters

| Method         | Window length (s) | Window step (s) | Number of windows | Audio length (s) | Number of bits |
|----------------|-------------------|-----------------|-------------------|-----------------|----------------|
| Reference      | 0.375             | 0.375           | 17                | 6.375           | 512            |
| Eigenvalue     | 0.03              | 0.02            | 108               | 2.17            | 512            |
| Wiener         | 0.03              | 0.02            | 108               | 2.17            | 512            |
| 2D DCT         | 0.03              | 0.02            | 128               | 2.57            | 512            |
Results
Conclusion

- We propose a combination of short overlapping windows and decorrelation of the energy spectrum of the audio signal for the generation of robust acoustic fingerprints.
- The required recording time is considerably reduced while increasing their robustness in conversational applications.
- The proposed fingerprints maintain the statistical properties that are necessary in an authentication process.
- The fingerprint generation methods do not add a significant computational load, and they can still be performed by low-power devices such as mobile phones.
Thank you for your attention