Seguretat en aplicacions web

GENERACIÓ DE CLAUS CRIPTOGRÀFIQUES AMB LA VEU

Autor: Òscar Raya i Casanova

Tutor: Antoni Martínez Ballesté

UAB Universitat Autònoma de Barcelona

UNIVERSITAT ROVIRA i VIRGILI

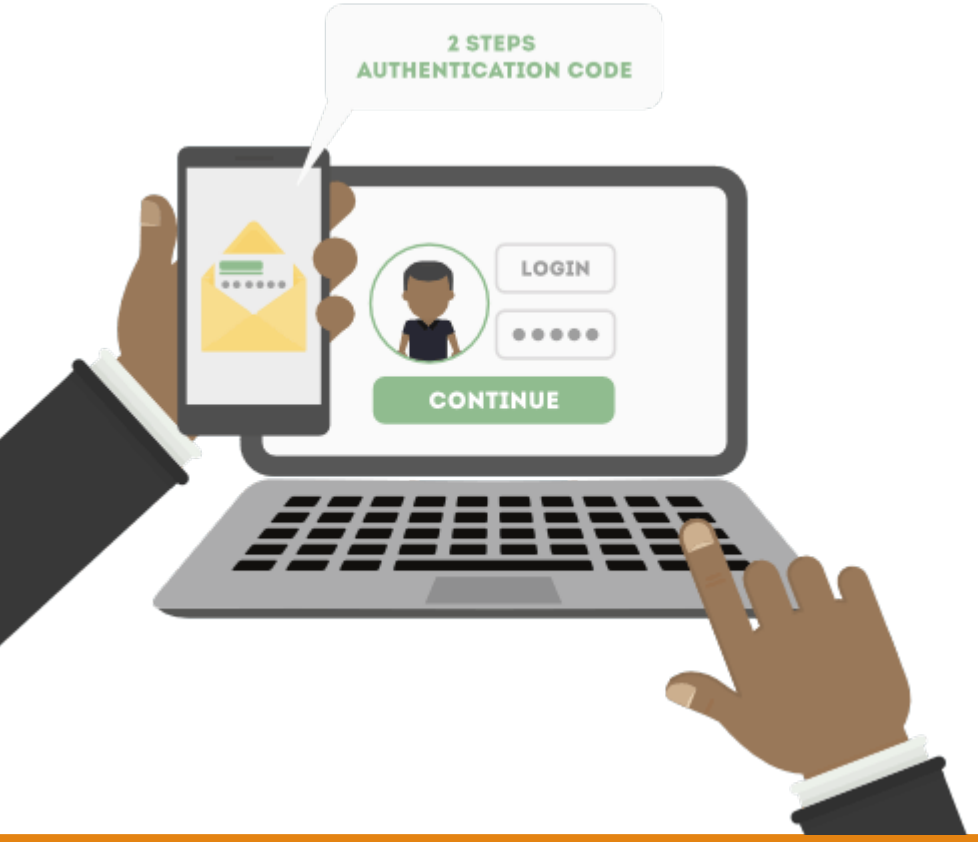
UOC Universitat Oberta de Catalunya

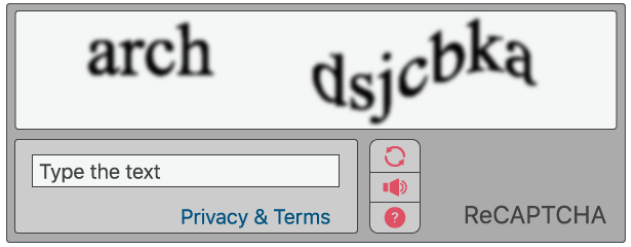
UIB Universitat de les Illes Balears

Índex

- Introducció
  - Objectius
  - Estat de l’art
- Metodologia de treball
- Anàlisi i disseny
- Implementació
  - Implementació MFCC
  - Implementació generador de Claus
  - Implementació aplicació Android
- Experiments i resultats
- Conclusions i treball futur
- Demostració

Introducció





Objectius

I. Estudiar quines característiques de la veu humana es poden utilitzar per identificar una persona.
II. Dissenyar un mètode robust amb el qual obtenir una contrasenya a partir de la veu.
III. Implementar una eina per posar en pràctica el mètode dissenyat.
IV. Provar el que s'ha implementat.

Estat de l’art

- Detecció de la veu o voice activity detection (VAD).
- Extracció de característiques:
  - Mel-frequency cepstral coefficients (MFCCs)
  - Linear predictive coefficients (LPCs)
  - Power normalized cepstral coeficients (PNCCs)
  - Spectral centroide frequency (SCF)
  - Cosine distance features (CDFs)
  - Gammatone frequency cepstral coefficients (GFCCs)
- Modelatge de la veu
  - Gaussian Mixture Model (GMM)
  - Probabilistic Neural Network (PNN)

Metodologia de treball

Anàlisi i disseny

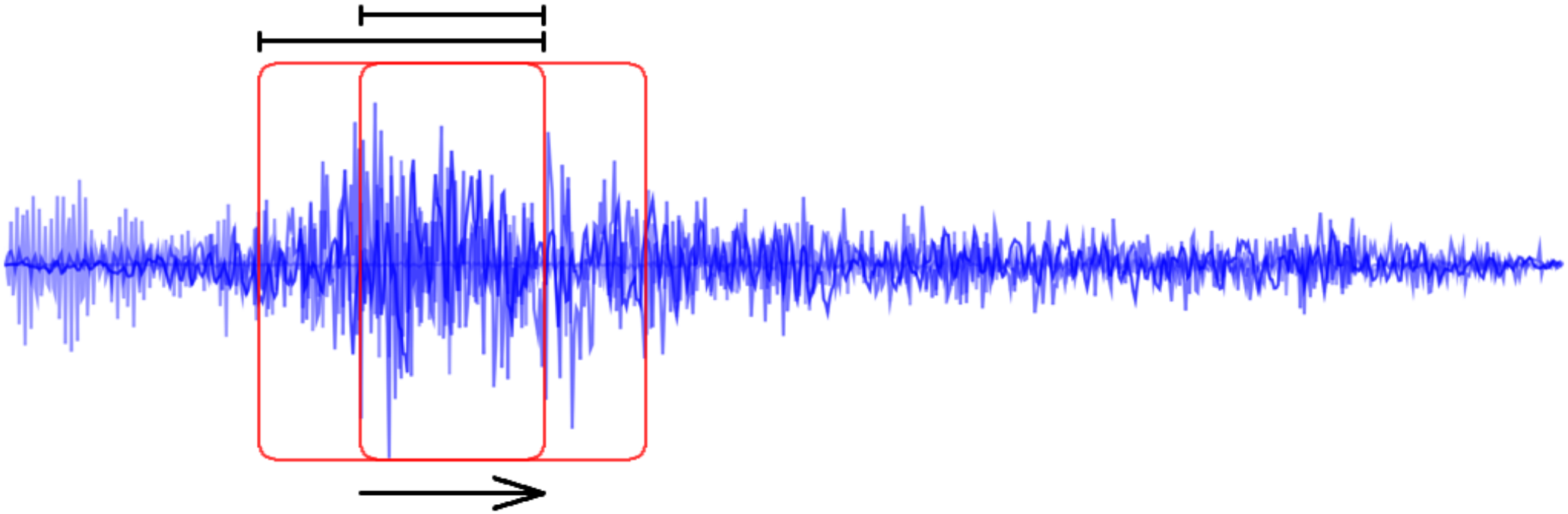
- Implementació d'un algorisme per detectar la presencia de veu (VAD). NO IMPLEMENTAT
- Implementació d'un algorisme que extregui característiques. Mes concretament s'extrauran els MFCCs.
- Implementació d'un algorisme per extreure una clau a partir de les característiques extretes. Aquest algorisme es basarà en els GMMs.
- Implementar una aplicació amb Android que sigui capaç de capturar la veu de l'usuari a traves del micròfon. S'ha implementat un mòdul de captura de la veu per tal de fer-lo servir en diverses pantalles.
- Establir una pantalla de registre a l'aplicació d'Android en la qual es registra la veu de l'usuari i se n'obté una clau diversos cops fins que es decideix quina clau ha d'utilitzar l'usuari (es busca utilitzar la clau que mes vegades s'hagi obtingut).
- Establir una pantalla d'inici de sessió a l'aplicació d'Android en la qual generi una clau a partir de a la veu de l'usuari.

Implementació MFCC

1. Dividir el senyal en trams petits i superposats.
2. Obtenir les potencies de densitat espectral (PSDs) de la transformada discreta de Fourier (DFT) de cada tram.
3. Aplicar filtres en l'escala de mel a les PSDs obtingudes.
4. Aplicar el logaritme en base 10 a les freqüències de mel.
5. Obtenir la transformada discreta del cosinus (DCT) dels logaritmes anteriors.

Els MFCCs són les amplituds de l’espectre obtingut amb la DCT. Normalment se solen utilitzar els coeficients del 2 al 13 i es descarten la resta. En el nostre cas s’ha decidit mantenir els coeficients del 2 al 14.

1. Divisió del senyal



2. Potencia de densitat espectral

DFT

$$X_n = \sum_{t=0}^{N-1} x_t e^{-knt}$$

$$k = \frac{2\pi i}{N}$$

PSD

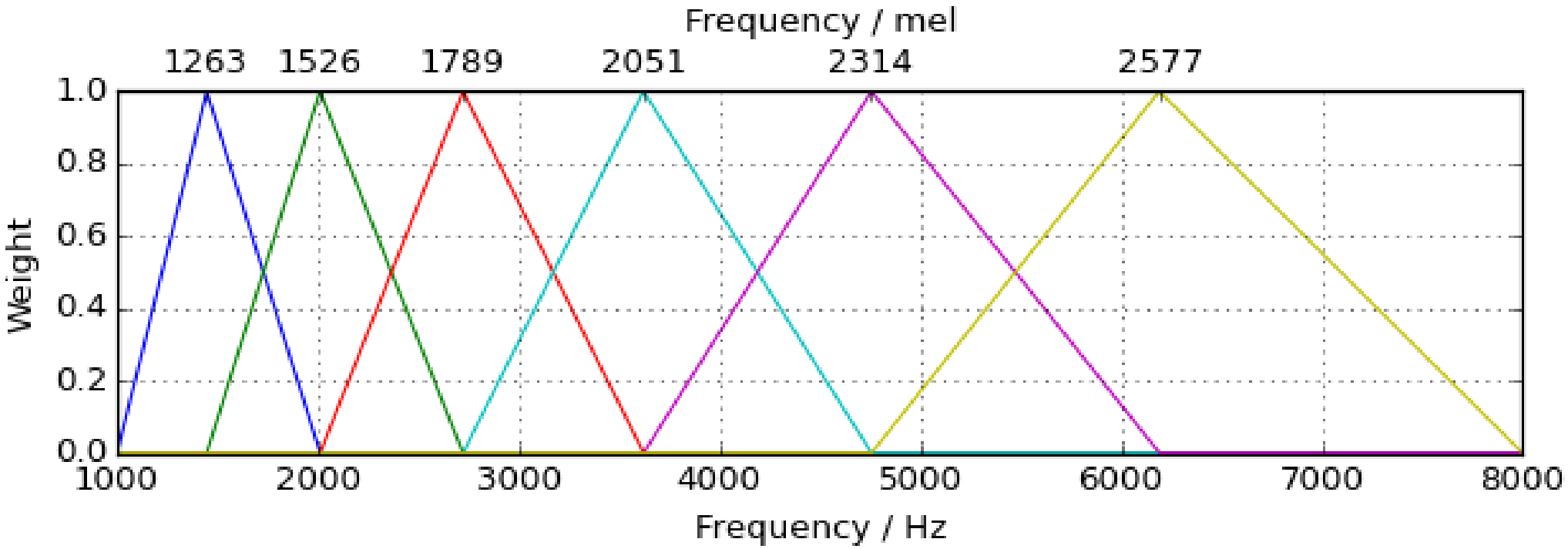
$$X_n = x_n \overline{x_n}$$

$$n = 0, \cdots, N-1$$

3. Filtres en l'escala de mel

$$M(f) = 2595\log_{10}(1 + f/700)$$

$$M^{-1}(f) = 700\left((m/2595)^{10} - 1\right)$$



4. Logaritme en base 10

Els humans no percebem la sonoritat de manera lineal.

SOLUCIÓ: Aplicar un logaritme a l'energia obtinguda dels filtres de mel.

*El nostre sistema utilitza un logaritme en base 10.

5. Transformada discreta del cosinus

$$F(k) = \frac{2c(k)}{N} \sum_{j=0}^{N-1} f(j) \cos\left[\frac{(j+0.5)k\pi}{N}\right]$$

$$c(k) = \begin{cases} 1/\sqrt{2}, & k = 0 \\ 1, & 0 < k < N \end{cases}$$

$$k = 0, \cdots, N-1$$

Implementació generador de claus

- Basat en GMMs (utilitzen distribucions normals).
- Utilitzem la mitjana i la desviació estàndard.
- S’assigna una lletra als diferents valors de mitjana i desviació estàndard i es concatenen.

Implementació aplicació Android

- S’ha implementat un mòdul per tal de capturar la veu en segon pla.
- S’ha introduït els mòduls necessaris per extreure els MFCCs i generar claus.
- L’aplicació consta de 4 pantalles:
  - Pantalla inicial – una pantalla simple amb dos botons per tal de poder registrar un usuari o iniciar sessió.
  - Registre d’usuari – una pantalla amb un boto que enregistra la veu quan es prem. L’usuari haurà de registrar la seva veu varis cops fins aconseguir obtenir una clau.
  - Inici de sessió – una pantalla amb un boto que enregistra la veu quan es prem. L’usuari registrarà la seva veu i l’aplicació n’obtindrà una clau la qual comprovarà. Si la clau és correcte es passa a la pantalla final.
  - Pantalla final – una pantalla que mostra que l’usuari ha aconseguit iniciar sessió.

Experiments i resultats

TIMIT data-set

Regió	Participants	Robustesa	Claus úniques	Claus compartides
1	49	16,33%	43	3
2	102	18,33%	61	19
3	102	17,94%	61	18
4	100	18,30%	63	13
5	98	18,67%	66	13
6	46	18,48%	38	4
7	100	17,30%	63	17
8	33	16,67%	24	4

Conclusions i treball futur

- El mètode no es prou robust ni sensible.
- Mitjana de 6 intens per poder accedir a l'aplicació (3 si s'utilitza la mateixa frase).
- Mitjana de 17 intents per registrar-se (9 si s'utilitza la mateixa frase).
- En un grup de 100 persones 40 d'elles compartiran clau amb algú altre.

- Treball futur:
  - Implementació VAD.
  - Millora algorisme DFT per tal que utilitzi fast Fourier transform FFT.
  - Afegir deltes i deltaDeltas als MFCCs.
  - Explorar l'ús d'altres característiques: LPCs, PNCCs, GFCCs, etc.
  - Afegir pesos diferents per a cada coeficient alhora de fer servir la funció de resum.

Demostració