Identity verification based on keystroke dynamics

Daniel Escobar Grisales
BSc. Electronic engineering student

Advisor: **Prof. Juan Rafael Orozco Arroyave Ph.D.**

Co-Advisor: **MSc. Juan Camilo Vázquez Correa**

Research group GITA, Antioquia University.

daniel.esobar@udea.edu.co

June 4, 2019

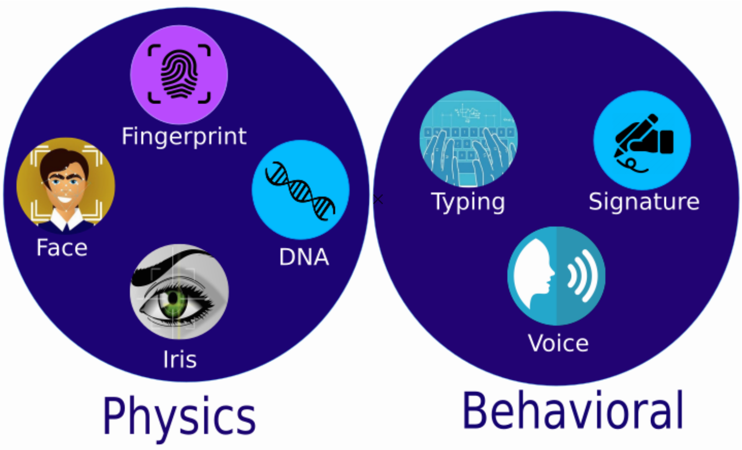
Outline

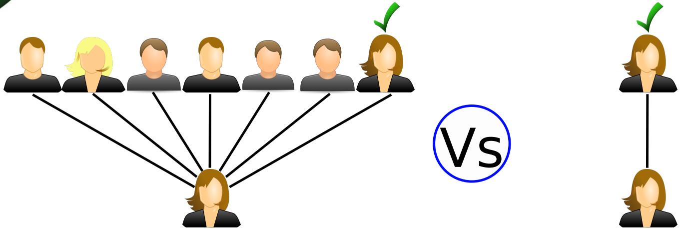
Introduction

Virtual education

Ingeni@
Soluciones TIC

Biometry


Fingerprint
Face
DNA
Iris
Physics
Typing
Signature
Voice
Behavioral



Identification Verification

UNIVERSIDAD DE ANTIOQUIA 1803

Objetives

Overall objective

- To develop a system for automatic registration and verification of identity of people based on their keystroke's dynamics.

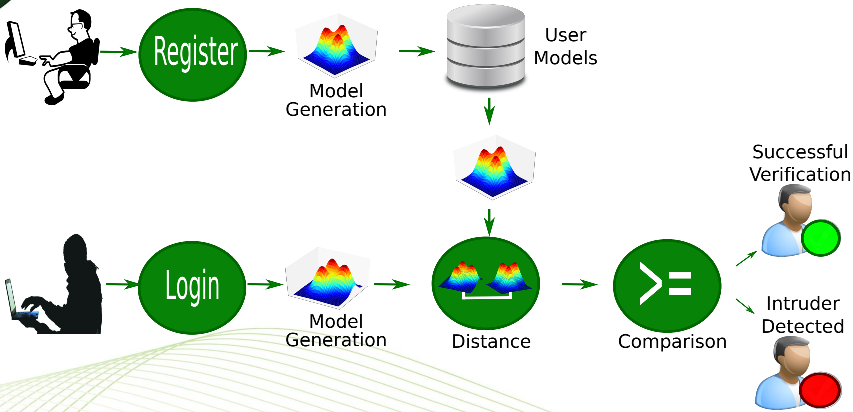
Specific objectives

- To implement characterization and identity verification algorithms based on keystroke's dynamics features.
- To evaluate:
  - FNR, False Negatives Rate.
  - FPR, False Positives Rate.
  - CUA, Cost to a User to Authenticate.
  - CUE, Cost to a User to Enroll.

UNIVERSIDAD DE ANTIOQUIA
1803

Methodology

Overall methodology

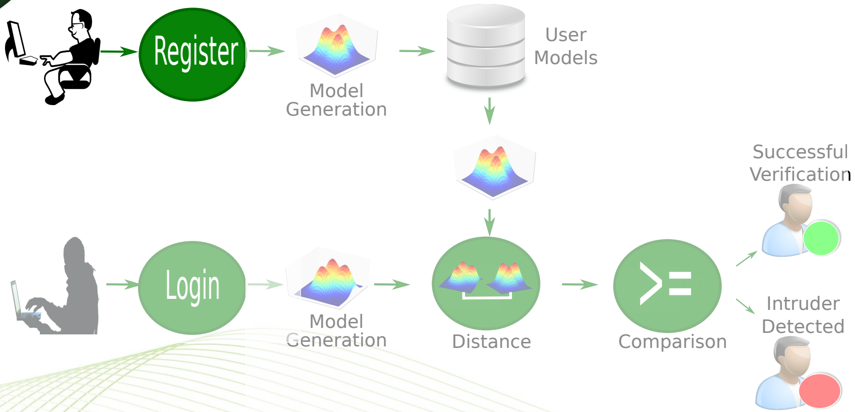


Database

Table: Participant's information

	Man	Woman
Number of subjects	116	54
Age ($\mu \pm \sigma$)	23.8 ± 5.8	24.4 ± 7.1
Bachelor students	102	44
Professional	6	7
Magisters	4	-
Doctors	4	3

Register


Register
Model Generation
User Models
Successful Verification
Login
Model Generation
Distance
Comparison
Intruder Detected

Register

UNIVERSIDAD DE ANTIOQUIA
1 8 0 3

Usuario:	Escriba su nombre
Cédula:	Escriba su cedula
Edad:	18
Género:	M
Escolaridad:	Estudiante de Pregrado
Nivel:	3
Programa:	Ingeniería Electrónica
Acepto términos:	

Ver términos y políticas de uso.

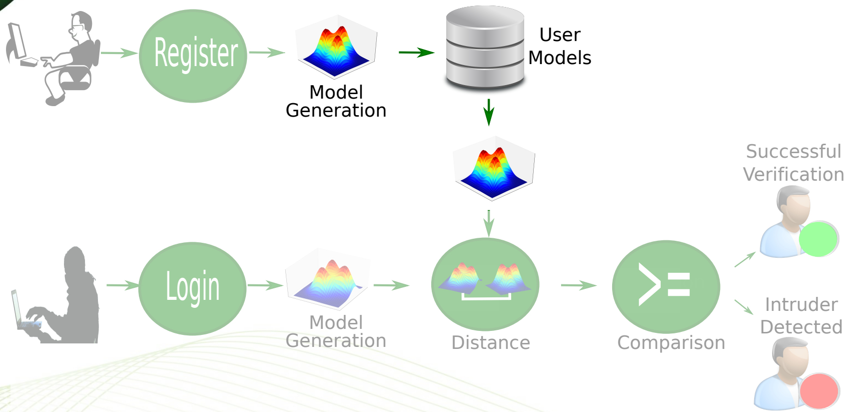
Registrarse

Escriba la siguiente frase, por favor incluya el punto al final de la frase:

El sapo de mi casa come kiwi, queso, zapallo y xoubas.

Limpiar Enviar

Model generation



Keystroke dynamic

Table: Example of data returned by the capture platform. word= "Hola", p: press, r: release.

Key	Code	Operation	Time(ms)
H	72	p	3301
o	111	p	3524
H	72	r	3556
o	111	r	3612
l	108	p	3644
l	108	r	3692
a	97	p	3716
a	97	r	3820

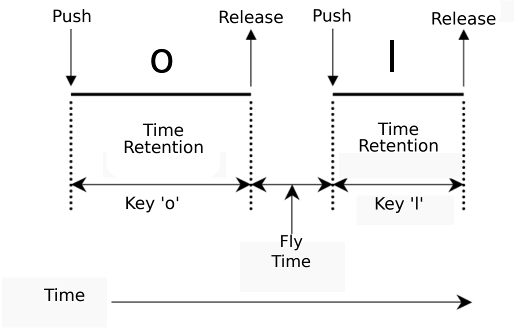
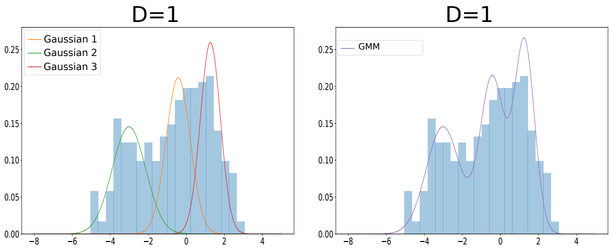


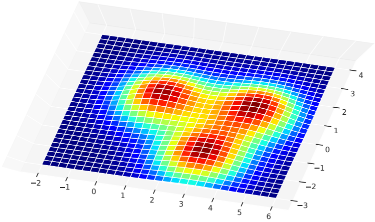
Figure: Retention time and fly time

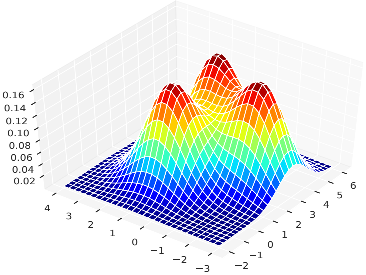
Gaussian Mixture Model (GMM)



$$f(x) = \sum_{m=1}^{M} \frac{c_m}{(2\pi)^{\frac{1}{2}} \sigma_m} \exp\left(-\frac{1}{2} \frac{(x-\mu_m)^2}{{\sigma_m}^2}\right)$$

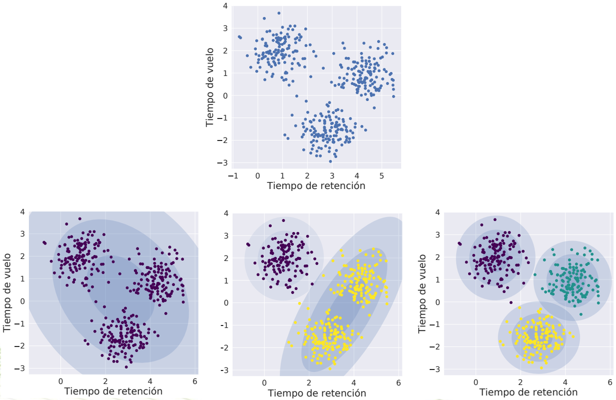
Gaussian Mixture Model (GMM)





$$f(\boldsymbol{x}) = \sum_{m=1}^{M} \frac{c_m}{(2\pi)^{D/2}|\boldsymbol{\Sigma}_m|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_m)^T \boldsymbol{\Sigma}_m^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_m)\right]$$

Gaussian Mixture Model (GMM)



Expectation Maximization algorithm

Algorithm:

1. First, initialize the parameters $\boldsymbol{\Theta} = \{c_m, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m\}$, for each Gaussian m.
2. Compute probability $h_m^{(j)}(t)$.
3. Change the parameters $\boldsymbol{\Theta}$ to maximize the previous probability.
4. Iterate steps 2 and 3 until convergence.

Step E:

$$h_m^{(j)}(t) = \frac{c_m^{(j)} \mathcal{N}(\mathbf{x}^{(t)}; \boldsymbol{\mu}_m^{(j)}, \boldsymbol{\Sigma}_m^{(j)})}{\sum_{i=1}^{M} c_i^{(j)} \mathcal{N}(\mathbf{x}^{(t)}; \boldsymbol{\mu}_i^{(j)}, \boldsymbol{\Sigma}_i^{(j)})}$$

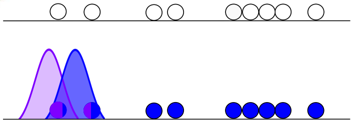
Step M:

$$c_m^{(j+1)} = \frac{1}{N} \sum_{t=1}^{N} h_m^{(j)}(t)$$

$$\boldsymbol{\mu}_m^{(j+1)} = \frac{\sum_{t=1}^{N} h_m^{(j)}(t) \mathbf{x}^{(t)}}{\sum_{t=1}^{N} h_m^{(j)}(t)}$$

$$\boldsymbol{\Sigma}_m^{(j+1)} = \frac{\sum_{t=1}^{N} h_m^{(j)}(t) [\mathbf{x}^{(t)} - \boldsymbol{\mu}_m^{(t)}][\mathbf{x}^{(t)} - \boldsymbol{\mu}_m^{(t)}]^{\mathsf{T}}}{\sum_{t=1}^{N} h_m^{(j)}(t)}$$

Expectation Maximization algorithm



Step E:

$$h_m^{(j)}(t) = \frac{c_m^{(j)}\mathcal{N}(\mathbf{x}^{(t)};\boldsymbol{\mu}_m^{(j)},\boldsymbol{\Sigma}_m^{(j)})}{\sum_{i=1}^{M} c_i^{(j)}\mathcal{N}(\mathbf{x}^{(t)};\boldsymbol{\mu}_i^{(j)},\boldsymbol{\Sigma}_i^{(j)})}$$

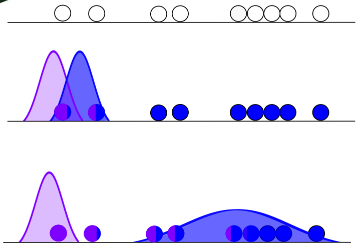
Step M:

$$c_m^{(j+1)} = \frac{1}{N}\sum_{t=1}^{N} h_m^{(j)}(t)$$

$$\boldsymbol{\mu}_m^{(j+1)} = \frac{\sum_{t=1}^{N} h_m^{(j)}(t)\mathbf{x}^{(t)}}{\sum_{t=1}^{N} h_m^{(j)}(t)}$$

$$\boldsymbol{\Sigma}_m^{(j+1)} = \frac{\sum_{t=1}^{N} h_m^{(j)}(t)[\mathbf{x}^{(t)}-\boldsymbol{\mu}_m^{(t)}][\mathbf{x}^{(t)}-\boldsymbol{\mu}_m^{(t)}]^{\mathsf{T}}}{\sum_{t=1}^{N} h_m^{(j)}(t)}$$

Expectation Maximization algorithm



Step E:

$$h_m^{(j)}(t) = \frac{c_m^{(j)}\mathcal{N}(\mathbf{x}^{(t)};\boldsymbol{\mu}_m^{(j)},\boldsymbol{\Sigma}_m^{(j)})}{\sum_{i=1}^{M} c_i^{(j)}\mathcal{N}(\mathbf{x}^{(t)};\boldsymbol{\mu}_i^{(j)},\boldsymbol{\Sigma}_i^{(j)})}$$

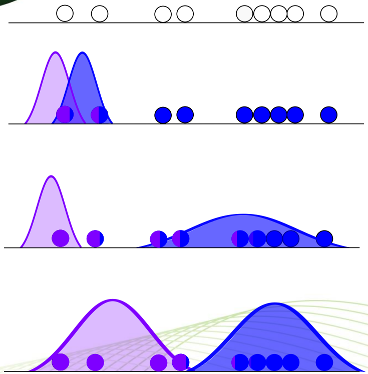
Step M:

$$c_m^{(j+1)} = \frac{1}{N}\sum_{t=1}^{N} h_m^{(j)}(t)$$

$$\boldsymbol{\mu}_m^{(j+1)} = \frac{\sum_{t=1}^{N} h_m^{(j)}(t)\mathbf{x}^{(t)}}{\sum_{t=1}^{N} h_m^{(j)}(t)}$$

$$\boldsymbol{\Sigma}_m^{(j+1)} = \frac{\sum_{t=1}^{N} h_m^{(j)}(t)[\mathbf{x}^{(t)}-\boldsymbol{\mu}_m^{(t)}][\mathbf{x}^{(t)}-\boldsymbol{\mu}_m^{(t)}]^{\mathsf{T}}}{\sum_{t=1}^{N} h_m^{(j)}(t)}$$

Expectation Maximization algorithm



Step E:

$$h_m^{(j)}(t) = \frac{c_m^{(j)}\mathcal{N}(\mathbf{x}^{(t)};\boldsymbol{\mu}_m^{(j)},\boldsymbol{\Sigma}_m^{(j)})}{\sum_{i=1}^{M} c_i^{(j)}\mathcal{N}(\mathbf{x}^{(t)};\boldsymbol{\mu}_i^{(j)},\boldsymbol{\Sigma}_i^{(j)})}$$

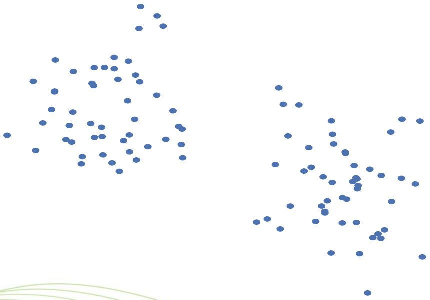
Step M:

$$c_m^{(j+1)} = \frac{1}{N}\sum_{t=1}^{N} h_m^{(j)}(t)$$

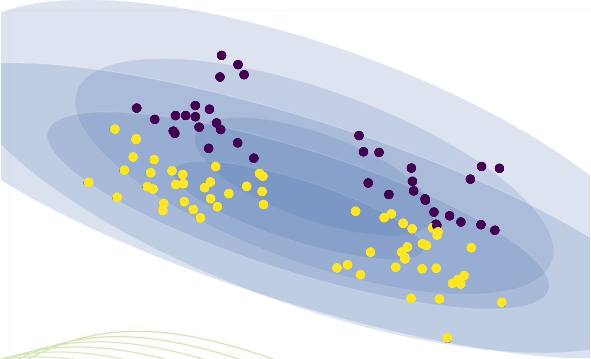
$$\boldsymbol{\mu}_m^{(j+1)} = \frac{\sum_{t=1}^{N} h_m^{(j)}(t)\mathbf{x}^{(t)}}{\sum_{t=1}^{N} h_m^{(j)}(t)}$$

$$\boldsymbol{\Sigma}_m^{(j+1)} = \frac{\sum_{t=1}^{N} h_m^{(j)}(t)[\mathbf{x}^{(t)}-\boldsymbol{\mu}_m^{(t)}][\mathbf{x}^{(t)}-\boldsymbol{\mu}_m^{(t)}]^{\mathsf{T}}}{\sum_{t=1}^{N} h_m^{(j)}(t)}$$

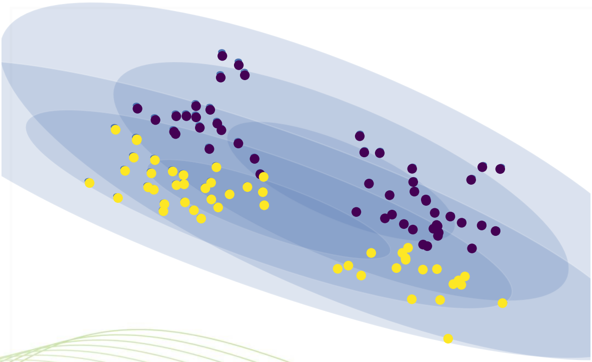
Expectation Maximization algorithm



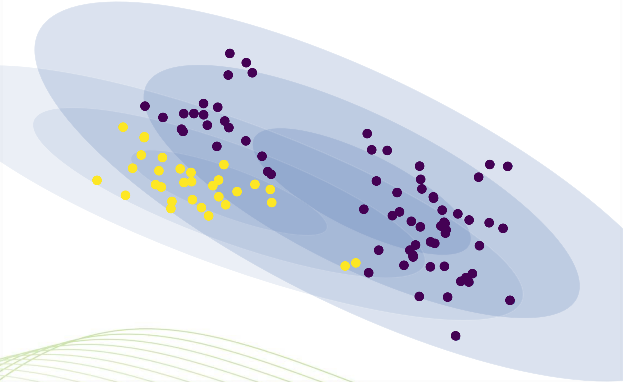
Expectation Maximization algorithm



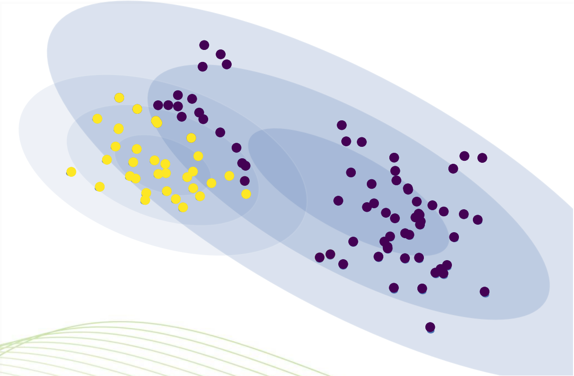
Expectation Maximization algorithm



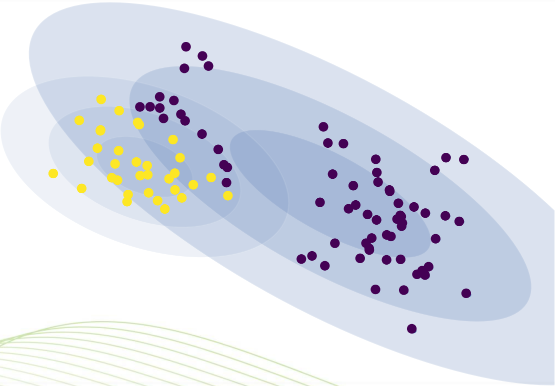
Expectation Maximization algorithm



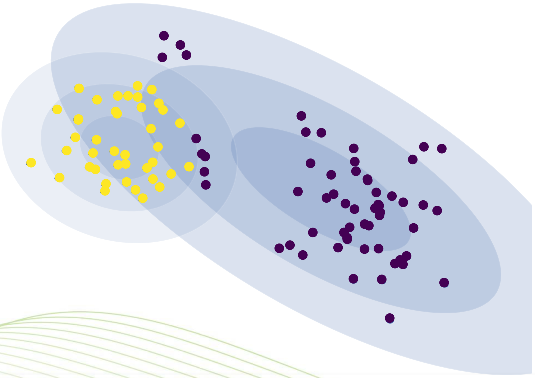
Expectation Maximization algorithm



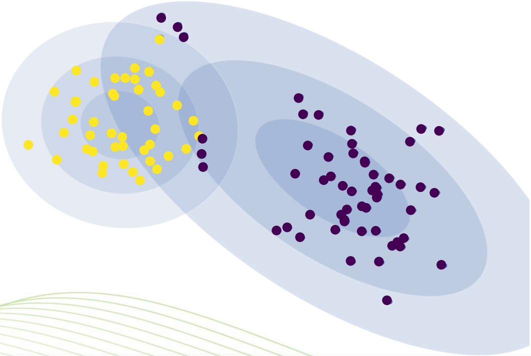
Expectation Maximization algorithm



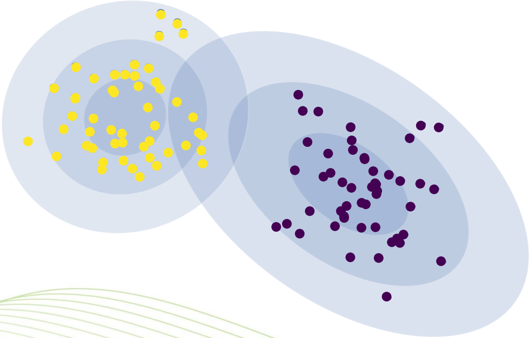
Expectation Maximization algorithm



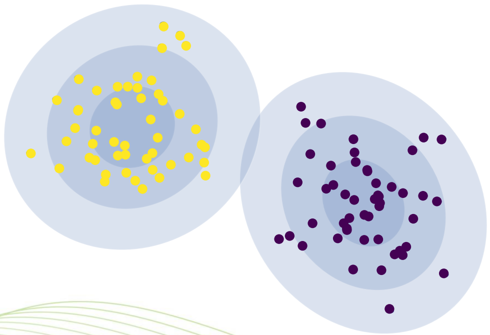
Expectation Maximization algorithm



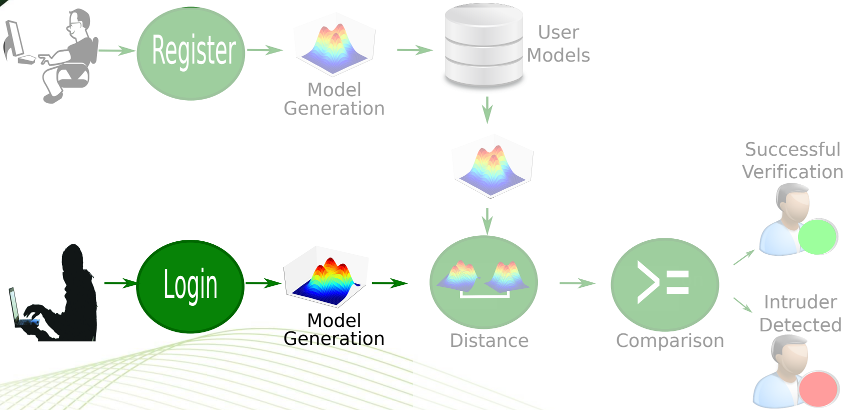
Expectation Maximization algorithm

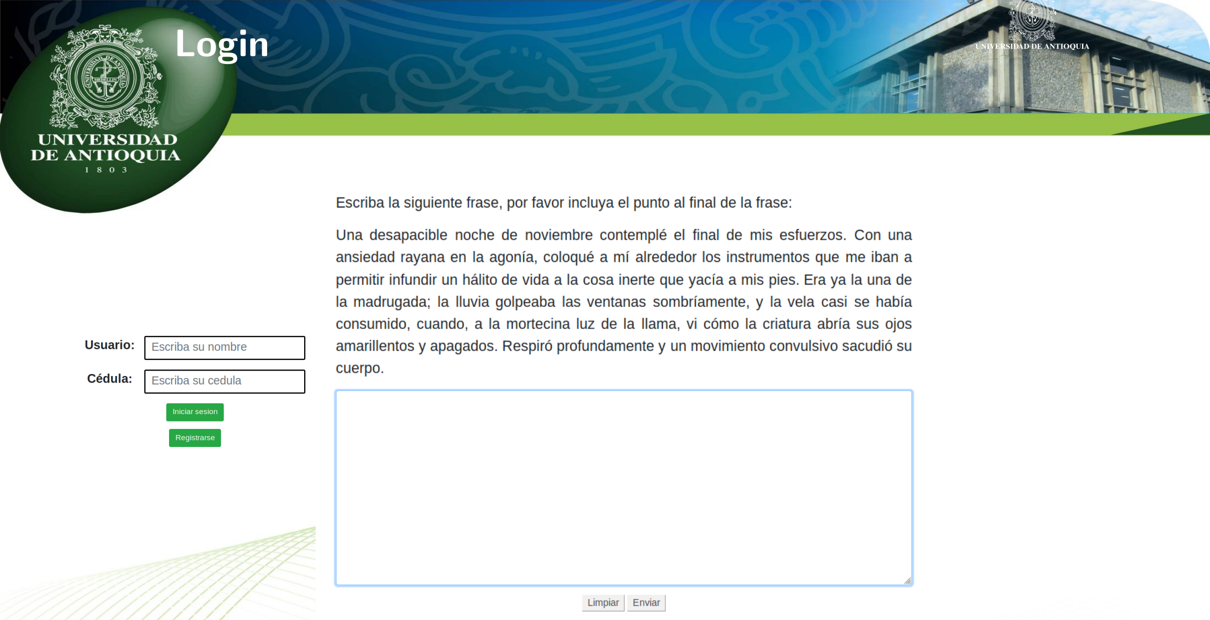


Expectation Maximization algorithm

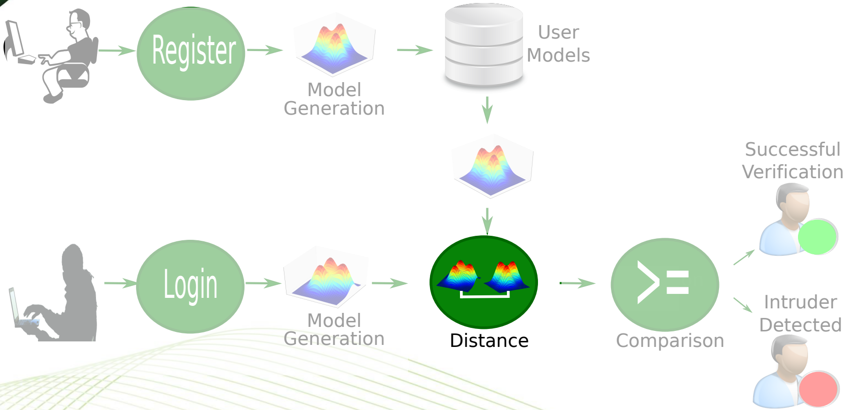


Login

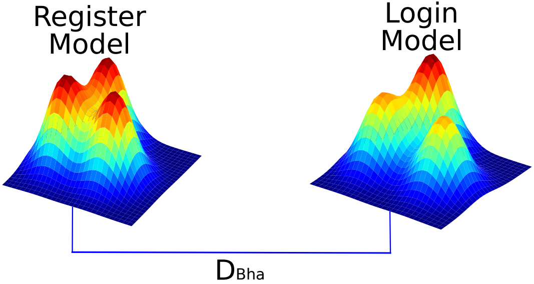

Register
Model
Generation
User
Models
Successful
Verification
Login
Model
Generation
Distance
Comparison
Intruder
Detected


Login
UNIVERSIDAD DE ANTIOQUIA
1 8 0 3
UNIVERSIDAD DE ANTIOQUIA
Escriba la siguiente frase, por favor incluya el punto al final de la frase:
Una desapacible noche de noviembre contemplé el final de mis esfuerzos. Con una ansiedad rayana en la agonía, coloqué a mí alrededor los instrumentos que me iban a permitir infundir un hálito de vida a la cosa inerte que yacía a mis pies. Era ya la una de la madrugada; la lluvia golpeaba las ventanas sombríamente, y la vela casi se había consumido, cuando, a la mortecina luz de la llama, vi cómo la criatura abría sus ojos amarillentos y apagados. Respiró profundamente y un movimiento convulsivo sacudió su cuerpo.
Usuario:
Escriba su nombre
Cédula:
Escriba su cedula
Iniciar sesion
Registrarse
Limpiar
Enviar

Distance


Register
Model Generation
User Models
Successful Verification
Login
Model Generation
Distance
Comparison
Intruder Detected

Distance

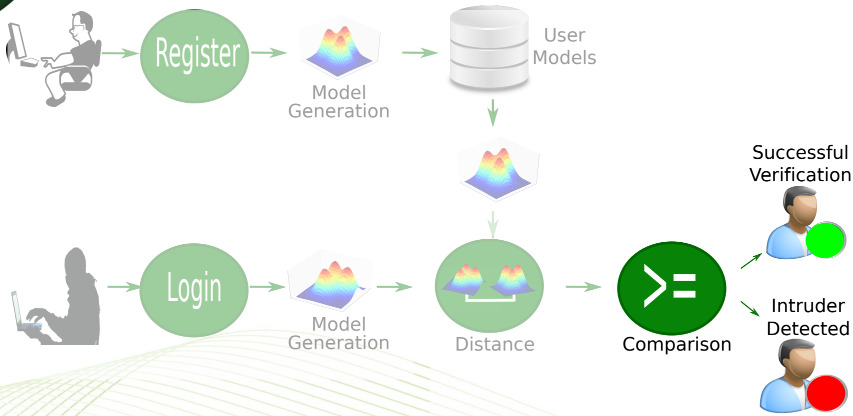


$$D_{Bha} = \mu_{Bha} + \Sigma_{Bha}$$

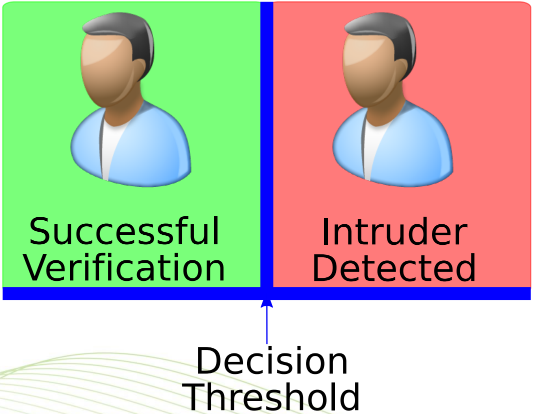
$$\mu_{Bha} = \frac{1}{8}\sum_{i=1}^{M}\left\{(\mu_{f_i} - \mu_{g_i})^{\intercal}\left[\frac{\Sigma_{f_i}+\Sigma_{g_i}}{2}\right]^{-1}(\mu_{g_i} - \mu_{f_i})\right\}$$

$$\Sigma_{Bha} = \frac{1}{2}\sum_{i=1}^{M}\left[\ln\frac{\frac{\Sigma_{f_i}+\Sigma_{g_i}}{2}}{\sqrt{|\Sigma_{f_i}||\Sigma_{g_i}|}}\right] - \omega_{Bha}$$

Comparison


Register
Model Generation
User Models
Login
Model Generation
Distance
Comparison
Successful Verification
Intruder Detected

Comparison


Successful
Verification
Intruder
Detected
Decision
Threshold

Performance and usability metrics

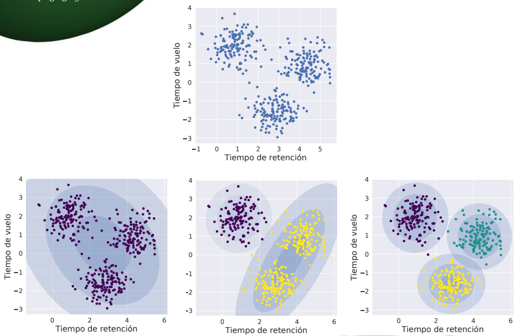
Performance metrics:

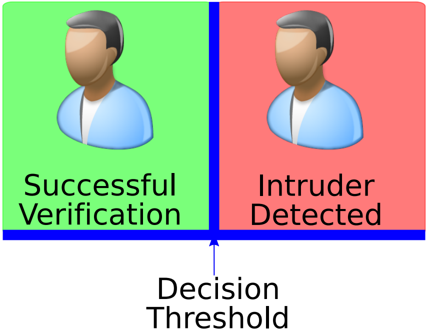
- FNR: Percentage of valid users that the system classified as intruders.
- FPR: Percentage of intruders that the system classified as valid users.

Usability metrics:

- CUE: Number of keystrokes that the user must make to register.
- CUA: Number of keystrokes that the user must make to authenticate.

Development stage





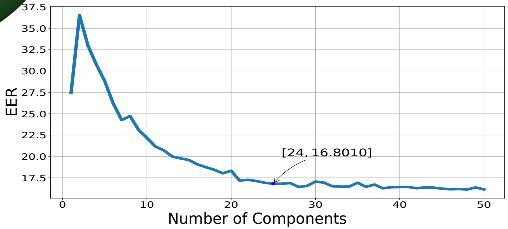
Number Components=?

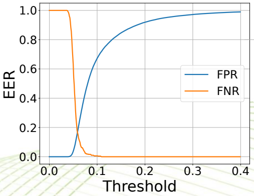
Decision Threshold=?

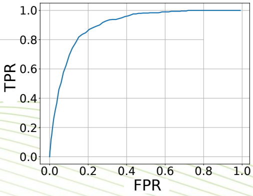
UNIVERSIDAD DE ANTIOQUIA 1803

Experiments and results

Development stage







Number Components	24
Threshold	0.059
FPR	16.02
FNR	16.92

Intrusive verification

Table: Performance and usability metrics, generating login models with known tasks by the register model.

Login Task	FPR	FNR	CUE	CUA
Task 1	18.68	10.00	314	54
Task 2	12.63	40.00	314	56
Task 3	**15.72**	**15.00**	**314**	**133**
Task 4	8.16	40.00	314	91

Valid access attempts: 20 (1 per user).
Intruder access attempts: 380 (19 per user).

No-intrusive verification (average distance)

Table: Performance and usability metrics, generating income models with unknown tasks by the register model, using an average distance.

Used Segment	FPR	FNR	CUE	CUA
2	13.66	17.14	314	204
3	13.69	12.86	314	308
4	12.21	11.43	314	416
5	**12.15**	**10.00**	**314**	**520**

Valid access attempts: 70 (1 per user).
Intruder access attempts: 4830 (69 per user).

Aplication

Usuario: Escriba su nombre

Cédula: Escriba su cedula

Iniciar sesion

Registrarse

Escriba la siguiente texto:

El cazador dudó si disparar al malvado lobo con su escopeta, pero luego pensó que era mejor usar su cuchillo de caza y abrir su panza, para ver a quién se había comido el bribón. Y así fue como con tan solo dos cortes logró sacar a Caperucita y a su abuelita, quienes aún estaban vivas en el interior del lobo.

Limpiar Enviar

Verificacion Exitosa

Terminar

Intruso Detectado

Terminar

http://pbiometriaconductual.udea.edu.co/

Conclutions

Conclutions

- From the analysis of the keystroke's dynamics, it is possible to verify a user's identity.
- This system doesn't need additional hardware.
- Acompanied by other biometric systems, it can help to reduce the rate of fraud.
- This methodology allows to verify the identity in two different modalities.

UNIVERSIDAD DE ANTIOQUIA
1 8 0 3

Future work

Future work

- Consider other methods to improve FPR and FNR, without increasing CUA.
- Try with different texts, in both modalities.
- Perform experiments with other languages.
- Classification of productivity and emotions with analysis of the keystroke's dynamics.