Anàlisi de robustesa basada en la identificació de patrons de teclat

Roger Muñoz

***Resum*—Una de les necessitats que sorgeixen en l'ús de sistemes d'aprenentatge virtual és la d'assegurar la identitat dels estudiants. Aquesta identificació es pot realitzar a través de sistemes biomètrics, com ara a través de la manera com els usuaris escriuen, és a dir, a través dels seus patrons de teclat. En aquest article es proposa el desenvolupament de dos models de patró de teclat de text lliure: un model bàsic basat en una gaussiana i un altre basat en mixtures gaussianes. Per tal d'avaluar aquests models, es realitzarà una comparativa dels models desenvolupats amb un producte comercial. Finalment es realitzarà un atac del sistema comercial a través de la generació de dades artificials obtingudes amb el model de mixtures gaussianes.**

***Index Terms*—Keystroke dynamics, identificació d'usuari biomètrica, atac al sistema comercial, model bàsic una gaussiana, model mixtures gaussianes**

I. Introducció

La generalització de l'accés a Internet i la popularització de les tecnologies digitals han permès que l'aprenentatge en línia sigui ja una realitat molt estesa, des de contextos més aviat informals (com ara aplicacions per a mòbil per aprendre idiomes), fins a l'educació superior formal (com ara les universitats totalment en línia).

En el context de l'educació reglada, una de les problemàtiques que sorgeixen és assegurar la identitat dels estudiants quan aquests realitzen les seves activitats dins d'un entorn d'aprenentatge virtual. Mentre que l'educació tradicional presencial habitualment resol aquesta problemàtica requerint als estudiants mostrar el document d'identitat quan aquests realitzen les proves d'avaluació, no existeix una solució senzilla en la versió virtual.

És precisament en aquest context en el qual l'anàlisi de les característiques biomètriques dels usuaris pot ser d'utilitat.

El terme biometria fa referència a la presa de mesures que descriuen característiques humanes (fisiològiques, de comportament, biològiques, etc.), que permeten descriure i diferenciar individus. La biometria es fa servir en el context de la identificació, per exemple, per reconèixer persones a partir de les empremtes digitals, fotografies o vídeos on hi apareix la cara, fragments d'àudio amb la veu, l'anàlisi forense de textos, i dels patrons de teclat a l'escriure en un dispositiu digital.

Aquest treball final de màster s'emmarca justament en l'anàlisi biomètric dels ritmes de patró de teclat dels usuaris dins del context de les activitats que aquests realitzen en sistemes d'aprenentatge en línia. L'objectiu principal d'aquest treball és doncs implementar i avaluar dos models d'anàlisi de patrons de teclat, i després fer-ne servir un d'ells per atacar-ne una versió comercial.

Aquest projecte ha estat inspirat en el projecte TeSLA (*Adaptive Trust-based System for Learning)*, un projecte europeu del programa marc H2020 que tenia com a objectiu principal proporcionar informació addicional de l'estudiant en activitats d'avaluació en línia. En la seva vesant tecnològica, TeSLA aportava una integració de diferents instruments amb l'objectiu d'oferir informació sobre l'autoria de les activitats dels estudiants.

En aquest article, en primer lloc es presenta l'estat de l'art pel que fa als sistemes biomètrics basats en el patró de teclat. Després, s'explica com es poden codificar les característiques biomètriques de patró de teclat per tal d'analitzar-les. A continuació, es presenten els dos models que s'han desenvolupat. Seguidament, es detallen els experiments que s'han realitzat per avaluar els models implementats. Un cop finalitzada l'explicació dels experiments, es presenten els resultats obtinguts, per realitzar una posterior anàlisi i discussió. Finalment, es realitza un atac intern a l'instrument comercial dins del context dels sistemes d'aprenentatge virtual.

II. Estat de l'art

Es poden trobar diferents tipus de sistemes biomètrics, segons la naturalesa de les característiques a analitzar. L'anàlisi facial [16], a través d'imatges o vídeos, permet extreure informació de l'usuari. Aquesta informació pot anar des de la identificació pròpia de l'usuari o el reconeixement del seu estat d'ànim [18], l'edat [7], etc. El reconeixement de veu permet reconèixer l'usuari [13], però també el contingut del que diu [9]. El reconeixement del tipus d'escriptura [3], és a dir, com un usuari escriu, també permet identificar l'usuari. I fins i tot el ritme de les pulsacions de com escriu també identifica l'usuari.

La identificació basada en patrons d'escriptura es pot aplicar en dues situacions diferenciades: en text fix i en text lliure. En el cas del **text fix**, de l'entrada analitzada se'n coneixen el conjunt de tecles, és a dir, l'entrada sempre és la mateixa i es coneix a priori. Un exemple seria l'anàlisi de com s'escriu una contrasenya [10]. En canvi, l'aplicació en el cas de **text lliure** analitza entrades diferents, textos lliures escrits per l'usuari, com ara el cos d'un correu electrònic o una resposta d'una pregunta oberta dins d'un entorn virtual d'aprenentatge.

En els paràgrafs següents es proporciona un resum molt breu dels articles de l'estat de l'art que cobreixen els dos contextos d'aplicació del reconeixement mitjançant patrons de teclat.

A la literatura es proposen diferents instruments de text fix. En aquest article [1], els autors van proposar una aplicació en text fix basada en una mesura estadística de proximitat.

En aquests articles [8], [5], es proposa l'ús de xarxes neuronals artificials amb el mateix objectiu. En canvi, en aquest article [6], es proposa l'ús de mixtures gaussianes i xarxes de creença profunda (que són un tipus de xarxes neuronals artificials).

A la literatura també es poden trobar diverses propostes per tractar textos lliures. En aquest resum [2], els autors analitzen implementacions de patrons de teclat de text lliure lliure mitjançant diferents algorismes. En aquest article [15], els autors van proposar un sistema basat en web per recopilar patrons de teclat dels navegadors.

A nivell d'anàlisi de patró de teclat, en aquest article [11] es resumeix que l'anàlisi dels temps en el ritme de teclat es pot utilitzar per identificar els usuaris quan utilitzen un sistema.

Un altre aspecte que es recull en la literatura és la utilització de dos tipus de models diferents [12]: generatius i discriminatius. El **model discriminatiu** intenta aprendre les fronteres entre les classes, mentre que el model generatiu aprèn la distribució de cada classe. El **model generatiu** permet, un cop entrenat el model, generar dades artificialment. És per aquest motiu que s'ha escollit el model generatiu per tal de poder dur a terme l'atac al model comercial, generant les mostres artificialment.

En aquest projecte s'implementaran dos models generatius emmarcats en l'aplicació de text lliure. Per una banda, la decisió d'utilitzar el model generatiu és perquè permetrà dur a terme l'atac al model comercial, tot generant les mostres artificialment. Per altra banda, en el context d'entorns d'aprenentatge virtual, les entrades dels usuaris no es coneixen i varien segons l'activitat i, per tant l'aplicació ha de ser de text lliure.

III. Codificació

El teclat és un dispositiu d'entrada basat en l'enviament de dos esdeveniments: *key down*, que es genera quan l'usuari prem una tecla, i *key up*, que es genera quan la deixa anar. Quan es parla de l'anàlisi de patró de teclat, el que realment s'analitza és el ritme (temps) com l'usuari prem i/o deixa anar les tecles del teclat. Això vol dir que s'analitzen els temps en els quals passen els esdeveniments *key down* i *key up*.

Depenent de quins esdeveniments i del número de tecles que es tinguin en compte, apareixen les següents definicions de temps (representades a la Figura 1):

Dwell És el temps que l'usuari té premuda una sola tecla key_1, és a dir, el temps entre l'esdeveniment key_1 *down* i key_1 *up*. No tothom està d'acord amb aquesta definició: alguns autors [6] prefereixen afegir dins del temps de *dwell* el temps de *flight*. En aquest projecte s'ha fet servir la primera definició.

Flight És el temps que l'usuari tarda en deixar anar una tecla key_1 i prémer la següent key_2, és a dir, el temps des de l'esdeveniment key_1 *up* i key_2 *down*.

Digraph És el temps que l'usuari tarda a prémer una tecla key_1 i prémer la següent key_2, és a dir, el temps des de l'esdeveniment key_1 *down* i key_2 *down*. Anàlogament, si en comptes de considerar dues tecles se n'analitzen tres o quatre, apareix el concepte de **trigraph** i **fourgraph**, respectivament.

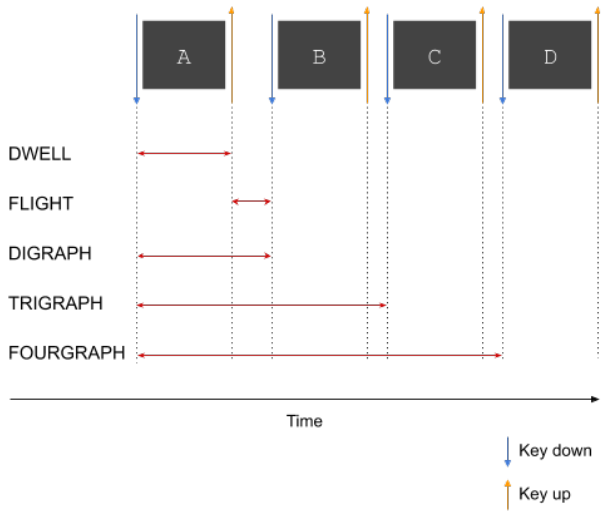


Figura 1: Definició dels temps codificats.

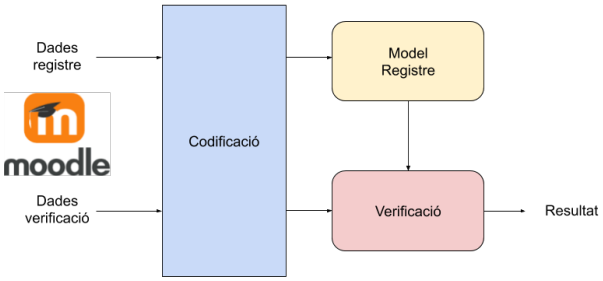


Figura 2: Procés de registre i verificació.

El model biomètric es pot construir doncs considerant el patró de teclat dels usuaris tenint en compte tots aquests temps per cada una de les tecles o combinacions de tecles que es puguin generar.

IV. Definició del model

La identificació basada en dades biomètriques es basa en la creació d'un model que representi les característiques de cada usuari. Així, en primer lloc s'entrena el model (és a dir, s'aprenen les característiques de l'usuari) i, després, es fa servir aquest model per obtenir un conjunt de probabilitats que decidiran la identificació de l'usuari. El primer procés s'anomena **registre** d'usuari i el segon **verificació**. A la Figura 2 es pot veure com l'usuari, a través del sistema d'aprenentatge virtual (en aquest cas el sistema Moodle), envia dades de registre que seran utilitzades per crear el model de l'usuari. Posteriorment, s'enviaran dades de verificació que seran analitzades i s'obtindrà un resultat d'identificació de l'usuari.

Formalment, el procés de registre és una funció F que a partir de les mostres de registre d'un usuari u, X_R^u, genera un model M^u d'aquell usuari:

$$M^u = F(X_R^u)$$

El procés de verificació té dues parts diferenciades. A la primera part, es fa servir el model entrat M^u per obtenir una

probabilitat P_{V_i} per cadascuna de les mostres de verificació X_{V_i}:

$$P_{V_i} = M^u(X_{V_i})$$

La segona part de la verificació parteix de les probabilitats obtingudes per cadascuna de les mostres i les combina per oferir un resultat final conjunt. Com que el context és d'una aplicació de text lliure, i per tant la llargària del text no està fixada, s'utilitza una puntuació, que s'incrementa o decrementa per cada probabilitat P_{V_i}, en funció de si la probabilitat indica que és l'usuari ($P_{V_i} > 0.7$) o no ($P_{V_i} \leq 0.7$), respectivament.

El valor amb el qual s'inicia la puntuació i els diferents increments i decrements són paràmetres que s'ajustaran a través de proves d'optimització de paràmetres. El valor a incrementar i decrementar per cada tipus de temps s'ajusta de manera individual.

Un detall a tenir en compte és que el model pot no saber calcular la probabilitat per alguna mostra d'entrada específica, X_{V_i}, per exemple, si aquesta mostra no s'ha observat durant el procés de registre (aprenentatge del model). En aquest cas, el model no pot decidir si és l'usuari o no, i la mostra és descartada.

A. Model bàsic

El primer model d'usuari que s'ha implementat és el model generatiu més simple: una distribució normal $\mathcal{N}(\mu, \sigma)$, on la mitjana (μ) i la desviació estàndard (σ) queden definides segons aquestes fórmules:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

on N és el número de mostres i x_i cadascuna de les mostres individuals ($i \in [1, N]$).

A partir del conjunt de mostres de registre, es calcula la mitjana i la desviació estàndard de cadascun dels temps presentats a la Secció III per cadascuna de les tecles presents al conjunt de dades.

A la Figura 3 es pot observar un conjunt de mostres (corresponents a una parella de tecles premudes per un usuari) i la gaussiana amb la qual han quedat modelades.

Per tal de poder crear el model de forma incremental, és a dir, sense haver d'esperar a tenir totes les mostres de registre de l'usuari per calcular-ne la seva mitjana i desviació, s'ha utilitzat l'algorisme naïve [17]. Aquest algorisme guarda, per cada tipus de mostra, el número total de mostres d'aquest tipus que s'han observat (N), la suma de temps (X) i la suma de quadrats del temps (XSQ). Quan arriba una nova mostra (t), aquests valors queden actualitzats segons les fórmules següents:

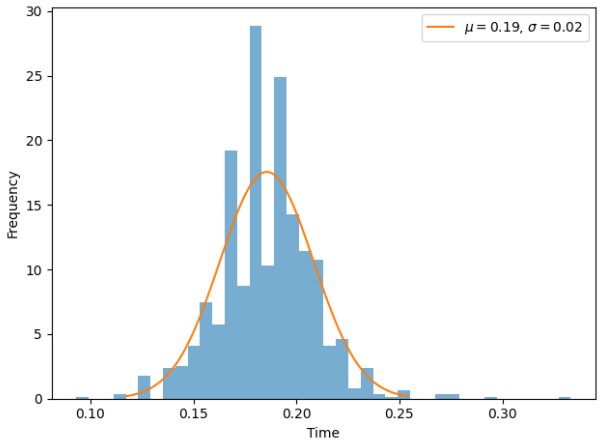


Figura 3: Exemple model bàsic basat en una sola gaussiana.

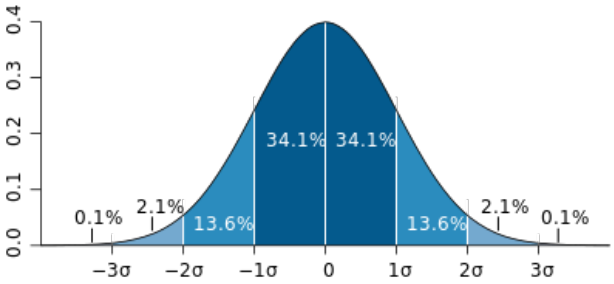


Figura 4: Distribució normal. Imatge original extreta de Wikimedia Commons.

$$N = N + 1$$
$$X = X + t$$
$$XSQ = XSQ + t^2$$

A partir dels valors N, X i XSQ d'un tipus de mostra es recuperen els valors de la mitjana i de la desviació estàndard amb les fórmules següents:

$$\mu = \frac{X}{N}$$

$$\sigma = \sqrt{\frac{XSQ}{N} - \mu^2}$$

En aquest model bàsic la probabilitat P_{V_i} és binària: si el temps rebut està a l'interval $[\mu - \sigma, \mu + \sigma]$, aleshores P_{V_i} és 1. En cas contrari, s'assigna P_{V_i} a 0. Aquest criteri s'ha aplicat ja que el 68.2% de les mostres d'una distribució normal (Figura 4) cauen dins d'aquesta àrea.

Aquest model s'ha implementat amb el llenguatge de programació *Python* i s'ha utilitzat la llibreria *numpy* per realitzar els càlculs.

B. Model mixtura de gaussianes

El segon model generatiu que s'ha implementat ha estat el model basant en mixtura de gaussianes (GMM, *gaussian*

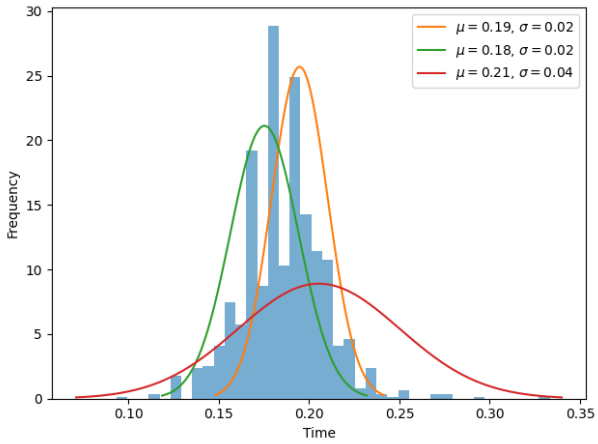


Figura 5: Exemple model mixtura de gaussianes

mixture model). Igual que en l'anterior, el model es genera a nivell de parella de tecles. En aquest cas, la probabilitat ve definida per la combinació de K gaussianes, cadascuna de les quals contribueix a la probabilitat amb pes π_k:

$$p(\boldsymbol{\theta}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{\mu_k}, \boldsymbol{\Sigma_k})$$

En la implementació realitzada, s'ha fixat el valor de K a 3, ja que amb aquest valor els resultats ja superaven el model anterior. A la Figura 5 es pot veure un conjunt de dades d'una parella de tecles d'un usuari en concret i les tres gaussianes amb les quals han quedat modelades amb el model GMM.

La principal diferència amb el model anterior és que a l'haver-hi més d'una gaussiana les dades s'ajusten millor al model. D'altra banda, l'augment de la complexitat d'aquest model fa necessari un temps de còmput més elevat a l'hora de realitzar el procés de registre de l'usuari. A més, aquest model no es pot generar incrementalment, és a dir, fins que no es tenen totes les mostres de l'usuari, no es pot crear el model de l'usuari.

Aquest model s'ha implementat amb el llenguatge de programació *Python* i s'ha utilitzat la llibreria *numpy* per realitzar els càlculs i *sklearn* per crear el model i avaluar-lo.

C. El model comercial

Aquest projecte inclou l'avaluació d'un tercer model de reconeixement de patró de teclat, que és un model comercial que es fa servir també en el context de l'aprenentatge en línia.

D'aquest model no se'n coneix la implementació concreta, i per tant, es tracta com a una **caixa negra** que proveeix de dues funcions bàsiques: el registre i la verificació.

En aquest cas, quan s'executa el registre d'un usuari, es crea un model M^u dins del sistema del proveïdor de l'instrument (model al qual no s'hi té accés), però que es pot fer servir posteriorment per a verificar l'usuari, i obtenir una probabilitat d'identificació.

V. Disseny d'experiments

En aquest apartat es descriuran els conjunts de dades utilitzats per a provar els models i, a continuació, el conjunt d'experiments que s'han realitzat per a avaluar-los.

A. Conjunt de dades

Els experiments que es descriuen en aquesta secció fan servir les dades recollides en tots els pilots de la institució UOC dins del marc del projecte TeSLA. Aquest conjunt de dades està anonimitzat i, a més, com que són dades biomètriques, aquestes no han sortit de la institució: el processament de les dades s'ha realitzat en dues màquines que es troben en el domini UOC.

Ha calgut tractar les dades abans de poder-les utilitzar en els experiments. El tractament que s'ha realitzat ha estat el següent:

1) Agrupar les dades per estudiants. Aquesta agrupació ha respectat la classificació de les dades de registre i verificació.
2) Descartar els estudiants que no tinguessin dades dels dos tipus: registre i verificació.

El conjunt de dades resultant és de 747 estudiants. Hi ha 72376 dades, de les quals 31301 són de registre i 41075 de verificació.

A part de la totalitat del conjunt de dades s'ha generat un subconjunt d'aquestes dades per tal de realitzar certs experiments amb més celeritat. El conjunt reduït té 11 estudiants que tenen dades de registre i verificació, amb 445 dades de registre i 316 de verificació.

L'estructura d'una mostra d'aquest conjunt de dades és en `JSON` que segueix aquesta estructura:

```
{
    "features": [
        {
            "type": integer,
            "code": string,
            "time": float
        },
        ...
        ]
}
```

Cada *feature* conté un *type* que indica de quin tipus de parella de caràcters és. La codificació d'aquest camp és la següent:

- 1 per *dwell*.
- 2 per *flight*.
- 3 per *digraph*.
- 4 per *trigraph*.
- 5 per *fourgraph*.

El camp *code* és un *hash* on hi ha codificada la informació de quines tecles operen en aquest *feature*. Utilitzant el *hash* en comptes d'enviar les tecles que s'estan utilitzant es realitza una ofuscació.

El camp *time* és el temps que ha tardat entre l'esdeveniment d'inici i l'esdeveniment final. Els esdeveniments concrets dependran del *type* que es representa en aquella *feature*.

L'ús del *hash* per representar tecles ve donada pel model comercial. Els models que s'han implementat (Secció IV) també s'ha utilitzat per tal de poder fer una comparativa més acurada amb el model comercial.

L'ofuscació de les dades proporcionada pel *hash* no aporta cap privadesa a nivell del missatge que s'envia ja que el domini és molt petit i constant per tots els usuaris (no es fa servir cap salt), de manera que un atacant pot fer un atac de força bruta amb un cost computacional molt baix.

B. Ajustament dels paràmetres

Tal com s'ha comentat a la Secció IV, per tal de fer servir el sistema s'han d'ajustar els paràmetres següents:

- La puntuació inicial per la qual es comença a comptar.
- El valor de l'increment i el decrement quan la mostra és de tipus *dwell*, *flight* o *digraph*.
- El valor de l'increment i el decrement quan la mostra és de tipus *trigraph*.
- El valor de l'increment i el decrement quan la mostra és de tipus *fourgraph*.

Per tal d'ajustar aquests paràmetres, s'han provat diverses combinacions de valors, avaluant l'exactitud del model resultant. El rang de valors que s'han provat en la puntuació inicial és $[0.1, 1]$, amb una granularitat de 0.1. En relació als paràmetres d'increment i decrement, inicialment s'han provat els mateixos valor que en la puntuació inicial, però s'ha observat que per valors més baixos s'obtenia millor resultat i, per tant, s'ha tornat a executar l'experiment amb un rang de $[0.01, 0.09]$ amb una granularitat de 0.01.

L'experiment consisteix en enviar primer les dades de registre de cada usuari i després les de verificació del mateix usuari. Llavors s'ha enviat les dades de verificació creuades, enviant per cada usuari registrat les dades de verificació corresponents a un altre usuari, seleccionat aleatòriament en cada execució.

Per tal d'escurçar els temps, aquest experiment s'ha realitzat amb el conjunt més petit de dades.

A partir dels resultats d'aquest experiment s'han fixat els paràmetres per realitzar els altres experiments.

C. Comparació dels models proposats amb una solució comercial

Una vegada fixats els set paràmetres de l'instrument, es procedeix a comprar els dos models proposats i la solució comercial. Per fer-ho, es procedeix de la mateixa manera que a l'experiment anterior: primer s'envien dades de registre per cada usuari, després dades de verificació corresponents a aquells usuaris, i finalment dades de verificació creuades.

Aquest experiment s'ha realitzat amb els dos models desenvolupats i la solució comercial.

Més enllà del resultat de la identificació, com a resultat addicional de l'experiment també s'ha guardat els temps d'anàlisi de cada mostra. En el cas dels models desenvolupats s'ha pogut calcular el temps real d'anàlisi, executant les proves en una màquina virtual amb quatre CPU's i vuit GB de RAM. En el cas del model comercial, els temps de registre no s'han pogut comparar amb els dos models implementats, ja que el sistema comercial utilitza un sistema de cues internes dins de l'instrument que n'alteren la seva mesura.

A diferència del primer experiment, en aquest s'ha fet servir el conjunt complet de dades.

D. Càlcul del mínim de dades necessàries de registre

Dins del context del projecte, el procés de registre d'un usuari en el sistema d'aprenentatge virtual cal que sigui el més ràpid, àgil i fàcil possible per l'usuari. Amb aquest experiment es vol comprovar quantes dades de registre d'usuari són necessàries per poder crear un bon model.

L'experiment consisteix en anar incrementant el número de mostres ($X \subseteq X_R^u$) de registre de l'usuari, des d'una sola dada fins a la totalitat del conjunt de registre disponible per cada usuari. Per cada mida del conjunt, s'ha seguit aquest algorisme:

- Enviar primer X mostres de registre d'usuari.
- Enviar totes les dades de verificació del mateix usuari.
- Enviar totes les dades de verificació creuades.
- Avaluar l'exactitud de l'experiment.
- Esborrar totes les dades del model de l'usuari.

En el cas del model comercial l'algorisme de l'experiment s'ha hagut de modificar perquè no permetia avaluar les mostres de verificació si no hi havia exactament 15 mostres de registre de l'usuari. Aquest és l'algorisme modificat:

- Enviar la mateixa mostra repetida $15 - |X|$ vegades.
- Enviar $|X|$ mostres de registre d'usuari. En aquest punt el registre de l'usuari és complet.
- Enviar totes les dades de verificació del mateix usuari.
- Enviar totes les dades de verificació creuades.
- Avaluar l'exactitud de l'experiment.
- Esborrar totes les dades del model de l'usuari.

El conjunt de dades utilitzat en aquest experiment ha estat el conjunt petit.

VI. Resultats dels experiments

Els experiments han estat avaluats amb les mètriques: ROC, AUC i exactitud.

La corba ROC (*Receiver Operating Characteristic*) és una representació gràfica de la sensibilitat d'un sistema respecte l'especificitat. Aquesta corba permet visualitzar com de bé classifiquen els models implementats i el comercial. La línia $y = x$ s'anomena línia de no discriminació, i correspondria a un classificador que generés prediccions aleatòriament.

El valor AUC (*Area Under Curve*) és l'àrea que queda sota la corba ROC. Com més pròxim sigui aquest valor a 1, millor és el classificador. Un classificador que donés respostes aleatòriament tindria un AUC de 0.5.

El valor d'exactitud (*accuracy*) es defineix com la capacitat d'un instrument de classificar correctament. Queda definida per la fórmula següent:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

on *TP* són els positius vertaders (*true postives*), *TN* són els negatius vertaders (*true negatives*), *FP* són els falsos positius (*false positives*) i *FN* són els falsos negatius (*false negatives*).

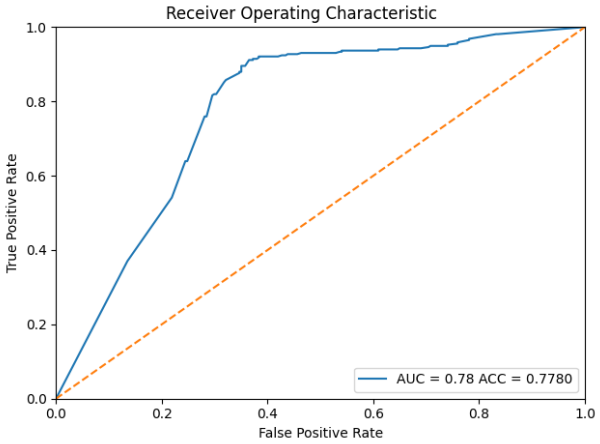


Figura 6: Resultat final de l'ajustament dels paràmetres del model bàsic.

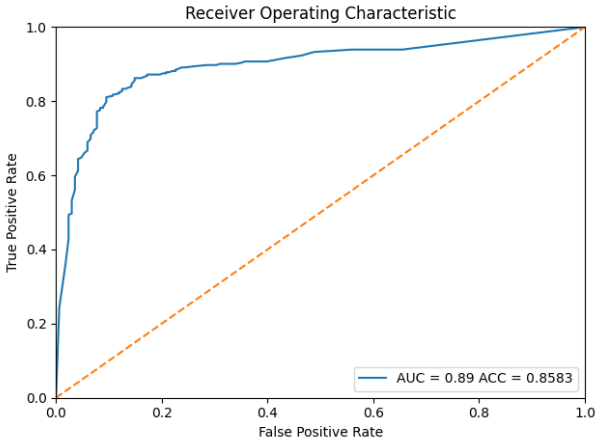


Figura 7: Resultat final de l'ajustament dels paràmetres del model de mixtures gaussianes.

Donat un model M^u que retorni 1 si les dades enviades corresponen a l'usuari u i 0 en cas contrari, aleshores els *TP* corresponen a les prediccions encertades en què $M^u = 1$, *TN* les prediccions encertades en què $M^u = 0$, *FP* les prediccions errònies en què $M^u = 1$ i, finalment, *FN* les prediccions errònies en què $M^u = 0$.

A. Ajustament dels paràmetres del model

La corba ROC dels models un cop ajustats els paràmetres es mostren a les Figures 6 i 7. Pel que fa el model bàsic el valor d'AUC obtingut és de 0.78 i l'exactitud 77.80%. En relació al model de mixtures gaussianes, l'AUC és 0.89 i el valor d'exactitud és del 85.83%.

Els paràmetres han quedat configurats de la següent manera:

- La puntuació inicial ha quedat fixada a 0.5.
- L'increment quan la mostra és de tipus *dwell*, *flight* o *digraph* ha quedat fixat a 0.02.
- El decrement quan la mostra és de tipus *dwell*, *flight* o *digraph* ha quedat fixat a 0.04.
- L'increment i el decrement quan la mostra és de tipus *trigraph* ha quedat fixat a 0.03.

Model	Temps mitjà registre	Temps mitjà verificació
Bàsic	0.3223	0.0221
Mixtures gaussianes	6.6686	0.3782
Comercial	0.0686	-

Taula I: Taula de temps mitjà de processament de mostres de registre i verificació (en segons).

- L'increment quan la mostra és de tipus *fourgraph* ha quedat fixat a 0.04.
- El decrement quan la mostra és de tipus *fourgraph* ha quedat fixat a 0.02.

B. Comparació dels models proposats amb una solució comercial

Els resultats d'aquest experiment es poden visualitzar a la Figura 8.

Es pot observar que el model bàsic (Figura 8a) és un classificador que està per sobre de la línia de no discriminació. El resultat AUC és de 0.71 i l'exactitud és de 67.67%.

El resultat del model de mixtures gaussianes (Figura 8c) és un classificador que es troba també per sobre la línia de no discriminació. El resultat AUC és de 0.84 i l'exactitud és de 79.33%.

El model comercial 8e és un classificador que, com els dos anteriors, està per sobre de la línia de no discriminació. El resultat AUC és de 0.88 i l'exactitud és de 80.22%.

El resultat de l'anàlisi de temps és el que es recull a la Taula I.

C. Càlcul de les dades de registre

Els resultats obtinguts per cada model són els que es poden veure a la Figura 8.

A la Figura 8b es pot observar que el model bàsic aconsegueix una exactitud acceptable a partir de 13 mostres, tot i que la desviació estàndard és molt gran. Això provoca que l'exactitud pugui arribar a caure per sota de 0.6, cosa que no és desitjable. Addicionalment, s'observa que no hi ha valor d'exactitud per 1 mostra, ja que el model no té suficients dades com per decidir si és l'usuari o no.

En el model de mixtures gaussianes s'observa (Figura 8d) un resultat bo ja que tenim una exactitud molt bona des del primer paquet de dades de registre. La desviació estàndard és petita per qualsevol valor del número de mostres.

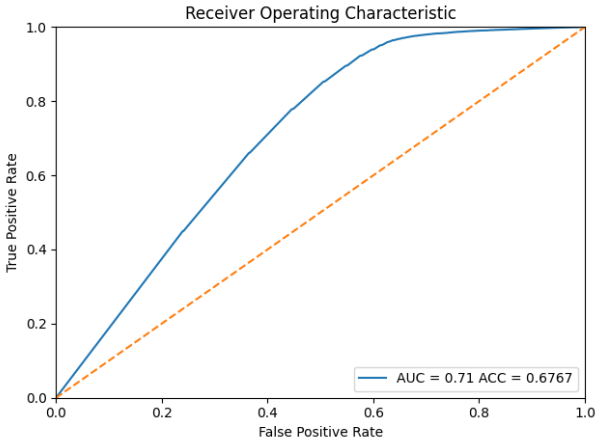
Pel que fa al model comercial, a la Figura 8f s'observa que a partir de la mostra 15 és on s'aconsegueix la millor exactitud. La desviació estàndard és molt gran.

VII. Anàlisi i discussió

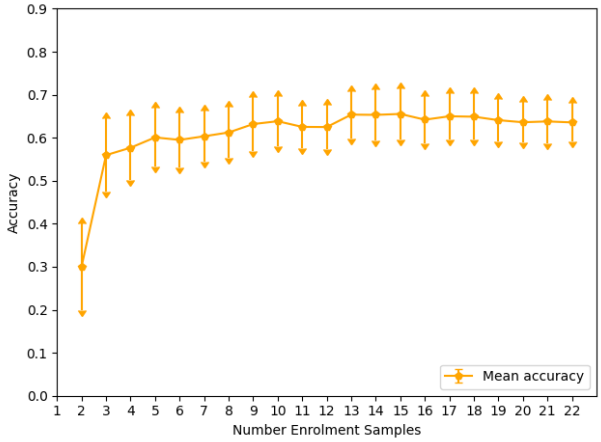
A continuació es detalla l'anàlisi i discussió per cadascun dels experiments.

A. Ajustament dels paràmetres del model

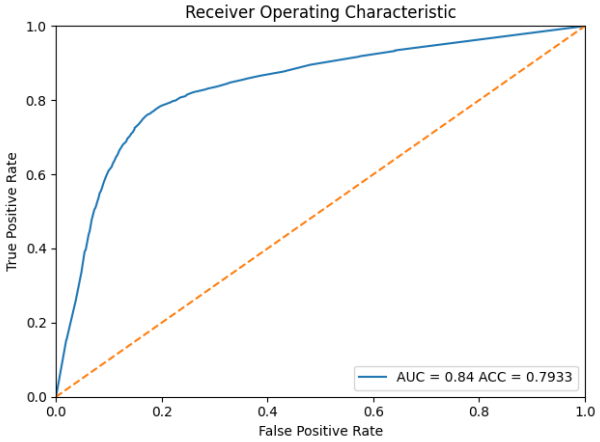
Amb aquest primer experiment ja es pot veure que el model de mixtures gaussianes dóna més bon resultat en relació amb el model bàsic, tal com s'esperava. Tant el valor d'AUC com



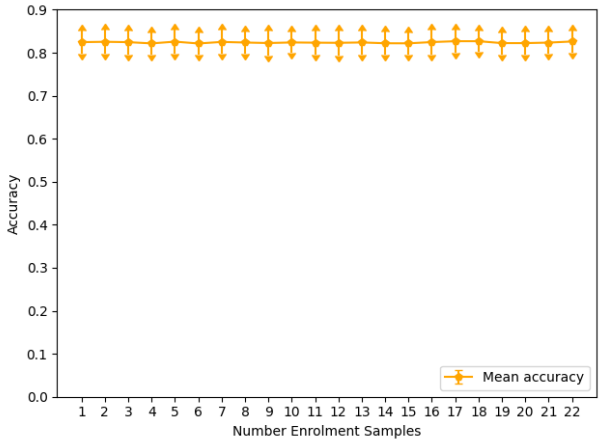
(a) Model bàsic: resultat amb tot el conjunt de mostres



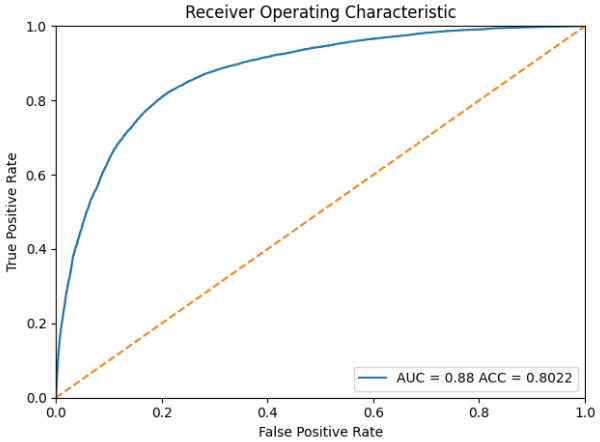
(b) Model bàsic: exactitud per nombre de mostres de registre



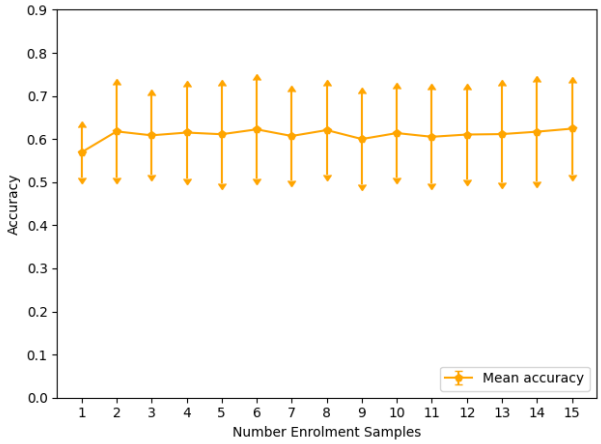
(c) Model mixtures de gaussianes: resultat amb tot el conjunt de mostres



(d) Model mixtura de gaussianes: exactitud per nombre de mostres de registre



(e) Model comercial: resultat amb tot el conjunt de mostres



(f) Model comercial: exactitud per nombre de mostres de registre

Figura 8: Corbes ROC i nivell mitjà d'exactitud segons el nombre de mostres de registre per als tres models comparats.

el d'exactitud del model de GMM són superiors als del model bàsic.

S'ha comprovat que el valor inicial en el qual s'inicialitza la puntuació no afecta al resultat obtingut de l'instrument.

Cal destacar que per tots els valors d'increments i decrements s'ha pogut observar que com més petits són més bons són també els resultats, és a dir, més alta és l'exactitud. Els valors seleccionats són els que han proporcionat una exactitud més alta.

B. Comparació dels models proposats amb una solució comercial

Es pot observar que el model bàsic 8a és el pitjor model ja que a valors baixos i mitjans de *true positive rate* té valors molt alts de *false positive rate*. També és el model amb pitjor valor AUC i exactitud.

En relació al resultat del model de mixtures gaussianes mostra millors resultats per valors baixos del *true positive rate*. El valor AUC està per sobre del model bàsic, però queda per sota del model comercial. El resultat és l'esperat: supera el model bàsic i s'aproxima al model comercial.

Pel que fa al model comercial, és l'instrument que millor resultat ha donat en aquest experiment amb una exactitud del 80.22% i un AUC de 0.88. Una possible explicació és que el model d'aquest instrument, a diferència dels altres dos, s'actualitza amb les noves mostres de verificació que li arriben de l'usuari, si es classifiquen com de l'usuari. Els resultats obtinguts en aquest instrument van en la mateixa línia que els resultats que es van publicar en el projecte TeSLA [4].

A nivell de temps d'execució dels models implementats (Taula I), el model més lent és el de mixtures gaussianes, i el més ràpid (pel que fa al registre), és el model comercial.

Dels resultats d'aquest experiment, es pot extreure que el millor model és el comercial, tot i que el model de mixtures gaussianes no s'allunya massa en relació a les mètriques d'exactitud i AUC. El model bàsic, com era d'esperar, és més ràpid que el GMM, però amb uns resultats bastant allunyats dels altres dos models.

C. Càlcul de les dades de registre

Com es mostra a les gràfiques amb els resultats (Figura 8), el model comercial és el model que necessita més dades de registre, 15. El segueix el model bàsic que amb 13 mostres en té prou. El model de mixtures gaussianes és el que necessita menys mostres: només amb 1 en té suficient per fer el registre d'usuari sense perdre exactitud.

Del resultat de l'instrument comercial, no s'esperava que amb 15 mostres de registre el valor d'exactitud fos tant baix, 60%, ja que en l'anterior experiment ha donat una exactitud d'un 80.22%. Donat aquest resultat, s'ha realitzat un altre experiment: enviar totes les dades del conjunt petit i treure'n la mitjana de l'exactitud. El resultat ha estat que amb aquest conjunt de dades l'exactitud és del 60% en comptes del 80%. Per tant, es pot afirmar que l'experiment s'ha realitzat correctament i que les diferències provenen del conjunt de dades analitzades.

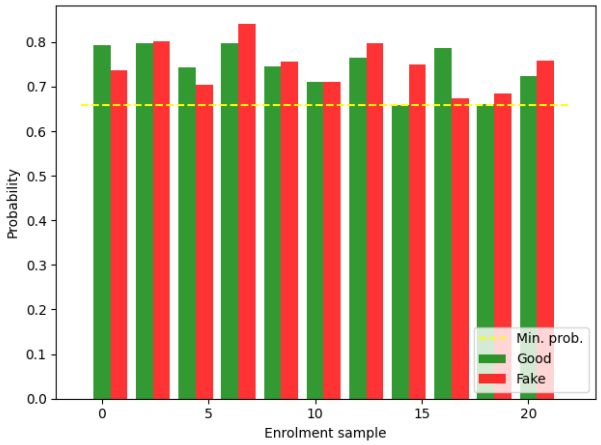


Figura 9: Atac del model comercial

VIII. Atac al model comercial

D'atacs als sistemes biomètrics n'hi ha de diverses naturaleses. Per exemple, un atacant de fora del sistema (és a dir, un atacant extern) pot intentar fer alguna acció fraudulenta, com ara la que es proposa en aquest article [14], on mitjançant l'anàlisi de temps de les pulsacions de teclat, entre altres, s'obté informació de les connexions xifrades en protocol *SSH*. Altres tipus d'atacs són realitzats per atacants interns, que són usuaris del propi sistema.

En aquest projecte es proposa fer un atac intern al sistema, és a dir, un atac on el propi usuari (que és un usuari legítim del sistema) intenta fer passar-se per un altre usuari.

L'atac que es vol dur a terme és un atac en temps de decisió: l'atacant no modifica les dades de registre dels usuaris, sinó que genera mostres artificials de verificació, imitant el seu comportament. En concret, l'atacant crea les dades fent servir el model generatiu de les mixtures gaussianes, i utilitza aquestes dades artificials per a atacar el model comercial.

L'atacant segueix l'algorisme següent:

1) L'atacant envia les seves dades de registre als dos models: mixtures gaussianes i comercial.
2) L'atacant genera dades (temps) de forma artificial amb el model de mixtures gaussianes que ha entrenat en el punt anterior.
3) Les dades són codificades per tal que les pugui entendre el sistema comercial.
4) Les dades codificades s'envien al sistema comercial.

Amb aquest atac, l'atacant és capaç de codificar un conjunt de dades, potencialment generades per qualsevol usuari, i fer que el model comercial les identifiqui com a vàlides per al propi usuari. En el context de l'aprenentatge en línia, això podria utilitzar-se, per exemple, per fer passar una activitat que ha realitzat un tercer com a pròpia.

Com es pot comprovar a la Figura 9, l'atac ha tingut èxit en tots els casos (el model ha considerat que les dades pertanyen a l'atacant amb probabilitat alta). Fins i tot, per a la majoria de mostres, les dades generades artificialment per a l'atacant han obtingut una probabilitat superior a les dades realment generades per l'usuari genuïnament.

Model	Exactitud	AUC	Núm. mostres registre	Temps mig registre	Temps mig verificació
Bàsic	67.67%	0.71	13	0.3223	**0.0221**
Mixtures gaussianes	79.33%	0.84	**1**	6.6686	0.3782
Comercial	**80.22%**	**0.88**	15	**0.0686**	-

Taula II: Taula de resum comparativa.

El resultat obtingut és força inquietant, ja que l'instrument comercial no ha mostrat cap tipus de resistència a l'atac.

IX. Conclusions

En aquest projecte s'ha implementat dos models de patró de teclat: un de bàsic amb una gaussiana i un de mixtures de gaussianes. Els paràmetres dels models s'han ajustat.

Els dos models implementats s'han comparat amb un instrument comercial, tant a nivell de temps necessari per a registrar els usuaris com dels resultats de la identificació (AUC i exactitud). A la Taula II es pot veure el resum dels experiments realitzats.

Finalment, s'ha realitzat amb èxit un atac intern al sistema comercial mitjançant la creació de mostres artificials amb el model de mixtures gaussianes. L'atac no ha sigut detectat per l'instrument comercial, que ha predit que les dades havien estat escrites per l'atacant en tots els casos.

X. Treball futur

Els resultats de l'avaluació dels models desenvolupats permet proposar fer-ne un prototipus. Durant aquest procés es podria optimitzar el model de dades de l'usuari que es genera. Per una banda, es podria reduir l'espai que ocupa en disc el model de dades de l'usuari, que és de l'ordre de 300KB i 800KB, per als models bàsic i de mixtures gaussianes, respectivament. En aquesta línia, s'hauria de fer una anàlisi de quines són les combinacions de tecles que permeten diferenciar més els usuaris. Per altra banda, es podria investigar com actualitzar el model de dades per tal que aquest segueixi sent vàlid al llarg del temps. En aquest projecte s'ha treballat amb models estàtics, és a dir, models que un cop generats no s'actualitzen mai.

Una altra de les línies d'investigació que s'obren en aquest treball és en relació als aspectes de privadesa de les dades d'usuari: es podria explorar si és possible poder fer anàlisis de patrons de teclat sense revelar el missatge que l'usuari està escrivint. En el context de sistemes d'aprenentatge en línia, això seria útil quan el proveïdor de l'instrument és diferent del proveïdor de la plataforma i, per tant, no té perquè tenir coneixement del missatge que escriu l'usuari.

Les dades utilitzades en els experiments van ser capturades per usuaris fent servir un teclat físic. Tenint en compte la gran variabilitat de dispositius que s'utilitzen per a accedir a les activitats en línia, seria també interessant comprovar si el model d'un usuari varia amb el dispositiu que fa servir o si, pel contrari, cal tenir un model diferent que representi a cada usuari fent servir cadascun dels dispositius dels què disposa.

Per últim, pel que fa a l'atac, la creació de models resistents a atacants interns és també una via de treball futur.

Agraïments

M'agradaria agrair la paciència i la dedicació del Dr. Xavier Baró en el projecte i tot el que l'envolta. M'agradaria agrair la utilització de les dades de patró de teclat de la institució UOC del projecte TeSLA. També m'agradaria agrair l'energia, l'empenta i el suport del meu entorn per tal que aquest projecte sigui una realitat.

Referències

[1] Mudhafar M Al-Jarrah. An anomaly detector for keystroke dynamics based on medians vector proximity. *Journal of Emerging Trends in Computing and Information Sciences*, 3(6):988–993, 2012.

[2] Arwa Alsultan and Kevin Warwick. Keystroke dynamics authentication: a survey of free-text methods. *International Journal of Computer Science Issues (IJCSI)*, 10(4):1, 2013.

[3] Miguel A Alvarez-Carmona, A Pastor López-Monroy, Manuel Montes-y Gómez, Luis Villasenor-Pineda, and Hugo Jair-Escalante. Inaoe's participation at pan'15: Author profiling task. *Working Notes Papers of the CLEF*, 2015.

[4] Xavier Baró, Roger Muñoz Bernaus, David Baneres, and Ana Elena Guerrero-Roldán. Biometric tools for learner identity in e-assessment. In *Engineering Data-Driven Adaptive Trust-based e-Assessment Systems*, pages 41–65. Springer, 2020.

[5] Sungzoon Cho, Chigeun Han, Dae Hee Han, and Hyung-Il Kim. Web-based keystroke dynamics identity verification using neural network. *Journal of Organizational Computing and Electronic Commerce*, 10:295 – 307, 2000.

[6] Yunbin Deng and Yu Zhong. Keystroke dynamics user authentication based on gaussian mixture model and deep belief nets. *International Scholarly Research Notices*, 2013, 2013.

[7] Xin Geng, Zhi-Hua Zhou, Yu Zhang, Gang Li, and Honghua Dai. Learning from facial aging patterns for automatic age estimation. In *Proceedings of the 14th ACM international conference on Multimedia*, pages 307–316, 2006.

[8] Erik Hellström. Feature learning with deep neural networks for keystroke biometrics: A study of supervised pre-training and autoencoders, 2018.

[9] K. . Lee, H. . Hon, and R. Reddy. An overview of the sphinx speech recognition system. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(1):35–45, 1990.

[10] Fabian Monrose, Michael K Reiter, and Susanne Wetzel. Password hardening based on keystroke dynamics. *International journal of Information security*, 1(2):69–83, 2002.

[11] Fabian Monrose and Aviel D Rubin. Keystroke dynamics as a biometric for authentication. *Future Generation computer systems*, 16(4):351–359, 2000.

[12] Andrew Y Ng and Michael I Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In *Advances in neural information processing systems*, pages 841–848, 2002.

[13] R. A. Rashid, N. H. Mahalin, M. A. Sarijari, and A. A. Abdul Aziz. Security system using biometric technology: Design and implementation of voice recognition system (vrs). In *2008 International Conference on Computer and Communication Engineering*, pages 898–902, 2008.

[14] Dawn Xiaodong Song, David A Wagner, and Xuqing Tian. Timing analysis of keystrokes and timing attacks on ssh. In *USENIX Security Symposium*, volume 2001, 2001.

[15] Alieksieiev Vasyl, Elena Sharapova, Olena Ivanova, Gorelov Denis, and Synytsia Yuliia. Web-based application to collect and analyze users data for keystroke biometric authentication. In *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, pages 917–922. IEEE, 2017.

[16] Yi-Qing Wang. An analysis of the viola-jones face detection algorithm. *Image Processing On Line*, 4:128–148, 2014.

[17] the free encyclopedia Wikipedia. *Algorithms for calculating variance*, 13 juny 2020 (accés 20 juny 2020).

[18] Amir Zadeh, Rowan Zellers, Eli Pincus, and Louis-Philippe Morency. Multimodal sentiment intensity analysis in videos: Facial gestures and verbal messages. *IEEE Intelligent Systems*, 31(6):82–88, 2016.