



# Exploración de correlaciones internas de los parámetros temporales generados en dinámicas de tecleo

Nahuel González<sup>1</sup>, Germán M. Concilio<sup>1</sup>,  
Jorge S. Ierache<sup>1</sup> , Enrique P. Calot<sup>1</sup> , y Waldo Hasperué<sup>2</sup>

<sup>1</sup> Laboratorio de Sistemas de Información Avanzados,  
Facultad de Ingeniería, Universidad de Buenos Aires,  
Buenos Aires, Argentina  
{ngonzalez,gconcilio,jierache,ecalot}@lsia.fi.uba.ar

<sup>2</sup> Instituto de Investigación en Informática (III-LIDI)  
Facultad de Informática Universidad Nacional de La Plata  
Investigador Asociado – Comisión de Investigaciones Científicas (CIC)  
whasperue@lidi.info.unlp.edu.ar

**Resumen** El análisis de cadencias de tecleo pertenece al ámbito de la autenticación biométrica por comportamiento. El comportamiento y el estado de ánimo de una persona modifican, sutil pero consistentemente, el patrón de escritura en un teclado de computadora y otros periféricos de entrada equivalentes. En este trabajo se exploraron, para un grupo de palabras elegidas del conjunto de datos de origen, las correlaciones internas entre los parámetros temporales, el rendimiento de clasificación en las tareas de identificación y verificación, y se buscaron grupos naturales de tipistas, entre los que se encuentran diestros y zurdos. Se observó que la influencia de cada parámetro temporal no es necesariamente local e inmediata, limitada a la tecla siguiente o una cercana, sino que pueden encontrarse correlaciones maximales entre letras con cuatro y hasta cinco grados de separación. Los resultados de el rendimiento en la clasificación, tanto en tareas de identificación como de autenticación, resultaron satisfactorios a pesar de la simplicidad del esquema utilizado, por lo que se propone la posibilidad de identificar el creador de un texto libre fragmentando la sesión por palabras y utilizando clasificadores individuales para cada una de ellas.

**Palabras claves:** keystroke dynamics, biometrics, identificación de usuario, autenticación de usuario, seguridad informática

## 1. Introducción

El análisis de cadencias de tecleo pertenece al ámbito de la biometría por comportamiento. El comportamiento y el estado de ánimo de una persona modifican, sutil pero consistentemente, su patrón de escritura en un teclado de

computadora y otros periféricos de entrada equivalentes. Por medio de estas variaciones individuales es factible verificar la identidad del usuario cuando esta es conocida o incluso identificarlo entre un conjunto de perfiles previamente entrenados con los que se cuente [9]. Lo que es más, ciertas características físicas como la mano dominante [14], o el estado emocional del usuario se pueden inferir de las mismas con precisión aceptable [8]; también se ha investigado la posibilidad de detectar tempranamente en las variaciones del patrón de escritura ciertas condiciones neurológicas de avance progresivo como el Alzheimer [17]. Utilizando únicamente la cadencia de tecleo de los usuarios se logró vincular múltiples cuentas o perfiles que difunden masivamente noticias falsas en Internet y las redes sociales durante la pandemia de COVID-19 [16]. El análisis de cadencias de tecleo en textos libres, con menos de veinte años de antigüedad [10], es un campo más complejo y menos explorado que el análisis de textos fijos como claves. Si bien algunos estudios tempranos han intentado zanjar la brecha entre la verificación de textos fijos y textos libres utilizando  $n$ -gramas [15], sólo se puede hablar de la madurez de la disciplina a partir de la introducción de distancias entre vectores de rango [1,10]. Desde el año 2013, apoyando al surgimiento de las investigaciones que articulan la detección del estado emocional con la dinámica de tecleo en textos libres, desde el Laboratorio de Sistemas de Información Avanzados de la Facultad de Ingeniería de la Universidad de Buenos Aires se comenzó a investigar en el área [11,5,3,4]. La utilización de contextos finitos para el modelado de cadencias de tecleo en textos libres ha sido propuesta en [11] con resultados promisorios en el ámbito de la verificación de identidad, como segundo factor de autenticación. En [7,6] se mostró la utilidad del método para identificar personas reales a partir del patrón de tipeo presentado en campos de texto libre en entornos web, en particular en el contexto de un e-commerce donde hay una gran variedad de tipos de dispositivos de entrada que afectan la efectividad final del método. En la Sección 2 se presenta el problema a investigar. El resto del trabajo se encuentra organizado de la siguiente manera: la Sección 3 describe la fuente de información que será minada en busca de hipótesis y la configuración del método aplicado, la Sección 4 discute los resultados obtenidos, y la Sección 5 presenta las conclusiones y futuras líneas de trabajo.

## 2. Descripción del problema

Si bien algunos periféricos de entrada como las pantallas táctiles de dispositivos móviles permiten obtener información de presión y aceleración, el enfoque clásico para análisis de cadencias de tecleo se ha centrado en los tiempos de escritura pues cualquier teclado convencional carece de medios para capturar parámetros adicionales. Los dos parámetros independientes que caracterizan la escritura en un teclado son el *tiempo de retención* (del inglés *hold time*), que es el lapso transcurrido entre la presión y la liberación de la tecla, y el *tiempo entre teclas* o *latencia* (del inglés *flight time*), que es el intervalo entre sucesivas presiones de tecla. Así, a cada carácter ingresado le corresponderían dos valores y la firma biométrica de un texto de  $n$  caracteres estaría conformado por un

vector de  $2n - 1$  tiempos; no  $2n$  pues el primer caracter carece de tiempo entre teclas. Se han utilizado también otros parámetros temporales, como por ejemplo el *tiempo de espera* que mide el intervalo entre la liberación de una tecla y la presión de la siguiente, o los  $n$ -gramas que agrupan las latencias de varias teclas adyacentes; todos estos, y muchos otros que han sido propuestos y evaluados en la literatura, derivan de los dos únicos parámetros independientes ya descritos.

En [11] se demostró que utilización de contextos finitos para el modelado de cadencias de tecleo aporta una mejora en la efectividad en el ámbito de la autenticación o verificación de identidad. En este trabajo se propone la posibilidad de identificar el creador de un texto libre fragmentando la sesión por palabras y utilizando clasificadores individuales para cada una de ellas. Se explorarán, para las palabras elegidas del conjunto de datos de origen, las correlaciones internas de parámetros temporales, el rendimiento de clasificación en las tareas de identificación y autenticación, y se buscarán grupos naturales de tipistas.

### 3. Experimentación

#### 3.1. Conjunto de datos

Para este trabajo se ha utilizado un conjunto de datos [13], subconjunto del utilizado en [12], que contiene multitud de sesiones de escritura en un teclado convencional de 158 usuarios, donde se han registrado las teclas presionadas y los tiempos de retención y latencia, junto con la identidad del escritor. Los tiempos se encuentran expresados en milisegundos, y por restricciones de la plataforma de captura en ciertas ocasiones se hallan redondeados al múltiplos de ocho más cercano, o algunos otros valores específicos. Las sesiones fueron capturadas en un entorno realista durante aproximadamente un año con usuarios de ambos sexos en un rango de edad entre 28 y 60 años, cuyas competencias mecanográficas variaban desde teclear con dos dedos hasta el manejo fluido. No hay motivos para suponer que la identidad etiquetada de los usuarios pueda contener errores, pues esta ha sido verificada al inicio de la sesión de escritura utilizando claves y/o otros factores de identificación biométrica. Los textos corresponden a lenguaje natural producido durante el transcurso de la labor cotidiana. Como las sesiones del conjunto de datos anterior varían en longitud y a los fines de construir un conjunto de datos tabular, se ha extraído de las sesiones los parámetros temporales correspondientes a un conjunto de cuatro palabras: QUE, DESDE, CONTROL, y EVOLUCIÓN. Estas palabras han sido elegidas por ser, dentro del conjunto de datos considerado, las que cuentan con mayor cantidad de observaciones para su extensión; las palabras más largas no son frecuentes y por lo tanto no brindarían suficientes observaciones para el análisis. De cada sesión en la cuál aparezca al menos una de las palabras indicadas se han extraído los tiempos de retención y latencia para cada una de las teclas que la conforman, y se ha etiquetado la fila con la identidad del usuario que las ingresó. En el Cuadro 1 se observan como ejemplo un grupo de filas con las columnas a los parámetros temporales de cada tecla de la palabra QUE. Las columnas denominadas HT son tiempo de retención, aquellas denominadas FT son latencias entre teclas

sucesivas, y el subíndice indica a qué tecla corresponde dicho tiempo medido en milisegundos. Todos los valores están expresados en milisegundos.

$HT_q$	$FT_q$	$HT_u$	$FT_u$	$HT_e$	$FT_e$	Usuario
$77 \pm 31$	$198 \pm 22$	$79 \pm 34$	$61 \pm 44$	$94 \pm 15$	$171 \pm 23$	585870
$72 \pm 1$	$156 \pm 78$	$88 \pm 16$	$251 \pm 54$	$68 \pm 8$	$184 \pm 71$	2085887
$83 \pm 3$	$311 \pm 224$	$67 \pm 3$	$121 \pm 20$	$70 \pm 5$	$86 \pm 30$	2038654
$141 \pm 1$	$257 \pm 8$	$86 \pm 16$	$156 \pm 0$	$117 \pm 8$	$149 \pm 7$	470698

Cuadro 1: Experimental Algunas filas para la palabra QUE

En el Cuadro 2 se resume la cantidad de atributos, instancias originales, y cantidad de instancias instancias filtradas de cada palabra del subconjunto de datos seleccionado para el análisis.

Palabra	Atributos	Instancias	Filtradas
QUE	3	5273	5112
DESDE	9	3322	3030
CONTROL	13	4965	4533
EVOLUCION	17	3200	2908

Cuadro 2: Resumen de los subconjuntos de datos

### 3.2. Preprocesamiento

El conjunto de datos de [12] fue preprocesado por medio de una herramienta propia para experimentos de cadencias de tecleo, a los fines de convertir el formato binario propietario en el cuál se encuentra almacenada la información original. En primer lugar, se buscó para cada largo de palabra aquella de mayor frecuencia, y luego se extrajo de cada sesión los tiempos de retención y latencia correspondientes a las elegidas. El resultado de este preprocesamiento fue almacenado en formato CSV para consumo de la herramienta de minería de datos. El archivo CSV resultante [13] incluye 49 columnas, de las cuáles 48 son de parámetros temporales y la restante etiqueta el usuario que los generó; la cantidad de filas asciende a 16760, y 158 usuarios. A los fines de simplificar la evaluación de subconjuntos de los datos por palabras, se produjeron conjuntos de datos reducidos en filas y columnas, para cada una de las palabras consideradas. De esta manera, con las 49 columnas originales se realizó una separación horizontal para generar cuatro conjuntos de datos derivados que representan los tiempos de escritura de las palabras QUE, DESDE, CONTROL, y EVOLUCION con 7, 11, 15, y 19 columnas respectivamente.

La herramienta de captura marca pausas en las sesiones, eliminando el valor de latencia entre teclas, toda vez que este excede un valor en torno a 1500

mseg., que depende parcialmente también de las medias móviles de los tiempos anteriores. Sin embargo, los valores usuales en la escritura se encuentran entre los 100 mseg. y los 500 mseg., en donde valores mayores suelen indicar pausas breves que no deben ser consideradas como parte de la cadencia natural. Para este fin se filtraron los valores extremos. Aproximadamente un 10 % de las instancias son excluidas por presentar valores extremos.

## 4. Resultados

### 4.1. Correlaciones internas en las palabras

Luego de filtrar los *outliers* se filtró el listado de correlaciones para comprender qué teclas se encuentran mayormente correlacionadas con otras dentro de la misma palabra. A continuación se muestran, para la palabra EVOLUCION, aquellos parámetros temporales cuyo coeficiente de correlación es mayor que 0,5. Este valor puede parecer escasamente significativo, pero debe tenerse en consideración el dominio particular del problema, la biometría comportamental, donde los parámetros son siempre muy ruidosos. Para distinguir la primera y segunda letra O de la palabra, en el Cuadro 3 se han utilizado subíndices,  $O_1$  y  $O_2$ .

Parámetro 1	Parámetro 2	Coef. correlación
$HT_I$	$HT_{O_2}$	0,693
$HT_V$	$HT_C$	0,681
$FT_E$	$FT_U$	0,591
$HT_E$	$HT_C$	0,551
$HT_{O_2}$	$HT_N$	0,549
$HT_{O_1}$	$HT_U$	0,526
$HT_I$	$HT_N$	0,524
$HT_E$	$HT_V$	0,510
$FT_C$	$FT_{O_2}$	0,507

Cuadro 3: Parámetros temporales con coeficiente de correlación mayor a 0,5 para la palabra EVOLUCION.

Se recorrió en orden decreciente la lista de correlaciones a los fines de descubrir, para cada parámetro temporal, con cuál otro manifiesta mayor correlación. El resultado para la palabra EVOLUCION puede observarse en la Figura 1, donde los nodos etiquetados con la letra corresponden a los tiempos de retención y los pequeños sin etiquetar a las latencias entre teclas. No es sorprendente que las mayores correlaciones se encuentren en los tiempos de retención pues estos son más estables, y relacionados con la velocidad promedio de escritura, que los tiempos de latencia entre teclas.

Destaca el hecho que la influencia de cada parámetro no es necesariamente local e inmediata, limitada a la tecla siguiente o una cercana, sino que pueden encontrarse correlaciones maximales entre letras con cuatro y hasta cinco grados

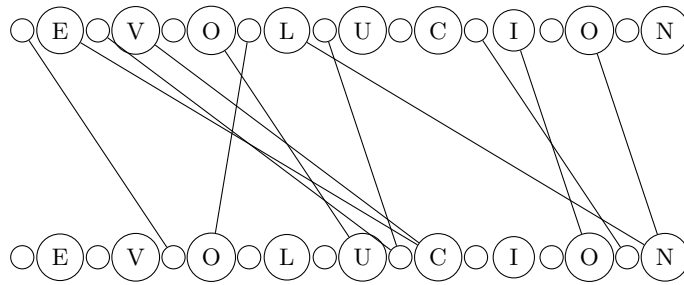


Figura 1: Pares de letras máximamente correlacionados

de separación como la E y la C, o la L y la N. Este fenómeno es relevante para el modelado de cadencias de tecleo utilizando contextos finitos [11] ya que al utilizar esta técnica el tamaño del contexto utilizado debe ser suficientemente amplio para capturar las influencias más significativas de las parámetros temporales de teclas anteriores si se desea obtener un buen rendimiento. La hipótesis de existencia de correlaciones fuertes entre teclas no adyacentes que revela el análisis del conjunto de datos agregado para todos los usuarios es particular también a los usuarios individuales al ser estos considerados por separado.

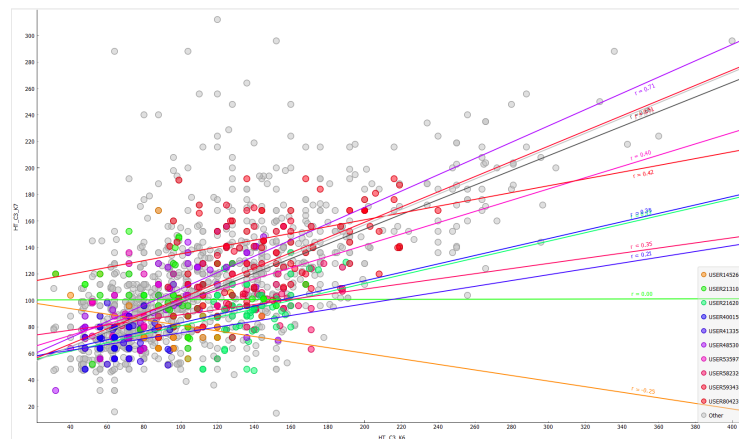


Figura 2: Detalle de correlación entre tiempos de retención de letras I y O en la palabra EVOLUCION

En la Figura 2 se observa el gráfico de dispersión de los tiempos de retención entre las teclas  $I$  y  $O_2$ , donde el eje  $y$  representa el tiempo de retención de la tecla  $I$  y el eje  $x$  el tiempo de retención de la tecla  $O_2$ , y muestra para cada usuario los valores de  $r$  que representan la correlación de las mediciones. Por ejemplo para el usuario 400150 presenta una correlación  $r$  de 0,71 aquellos que presentan

mayor correlación. Casi todos los usuarios presentan similar tendencia, aunque no debe descartarse la existencia de contraejemplo como el usuario 145264 con  $r = 0,25$ .

#### 4.2. Rendimiento de clasificación en las tareas de identificación y autenticación

El análisis clásico de cadencias de tecleo en textos libres ha utilizado indiscriminadamente la información disponible en la sesión, por ejemplo a través de métricas de desorden [1,10]. La ponderación de distintos contextos en base a su capacidad de discriminar ha demostrado mejorar el rendimiento en la clasificación [11], sobre todo al utilizar la fragmentación de las sesiones en fronteras no arbitrarias, sino utilizando inicio y fin de palabras. Aquí se explorará brevemente el rendimiento de clasificación en las tareas de identificación y autenticación en palabras individuales. Para tal fin se compararon cuatro clasificadores (kNN, Random Forest, SVM, y redes neuronales) tanto en una tarea de identificación como de autenticación. En la primera de ellas se intenta asignar la cadencia de tecleo al usuario correspondiente entre todos los que aparecen en el dataset, mientras que en la segunda se trata de un problema de clasificación binaria donde se utilizan tantas instancias del usuario legítimo como de otros usuarios, estos últimos marcados como impostores. Para la primera, de mayor dificultad, es esperable un rendimiento inferior. En el Cuadro 4 pueden observarse los resultados, junto con los intervalos de confianza para asegurar su significación estadística, para las palabras QUE, DESDE, CONTROL y EVOLUCION. Nótese que la diferencia de magnitud en los intervalos para ambos tipos de tareas se explican por la diferencia en la cantidad de filas; mientras que la tarea de identificación utiliza todas las filas del dataset para cada palabra, en el orden de varios miles, la tarea de autenticación utiliza únicamente el doble de la cantidad del usuario evaluado (las del usuario más idéntica cantidad de impostores), en el orden de pocos cientos. No es un error que el intervalo de confianza para las tareas de identificación, que ha sido verificado, sea el mismo en todas las filas. A primera vista el rendimiento en la clasificación es sorprendentemente bueno para las palabras CONTROL y EVOLUCION, pero no debe olvidarse que las palabras elegidas son las más comunes del dataset [13] para el largo dado. Las palabras escritas con mayor frecuencia estabilizan la cadencia de tecleo, en contraste con aquellas más raras que suelen ser más ruidosas y por lo tanto no exhibirían métricas de precisión tan optimistas. Sin embargo, estos resultados son lo suficientemente alentadores como para considerar la posibilidad de identificar el creador de un texto libre fragmentando la sesión por palabras y utilizando clasificadores individuales para cada una de ellas. Por otro lado, también se puede observar que con palabras muy cortas como QUE o DESDE, resulta imposible identificar unívocamente al usuario. Sin embargo se mantienen resultados satisfactorios en la determinación de autenticación de usuarios, en este orden fortalece el objetivo primario de la investigación presentada.

Palabra	Modelo	Precisión identif.	Precisión autenticación
QUE	kNN	19,3 ± 1,06	72,9 ± 1,28
	SVM	26,3 ± 1,18	74,2 ± 1,26
	Red neuronal	27,3 ± 1,20	68,9 ± 1,33
	Random Forest	25,0 ± 1,17	71,6 ± 1,30
DESDE	kNN	35,3 ± 1,71	83,0 ± 1,38
	SVM	42,1 ± 1,75	93,2 ± 0,95
	Red neuronal	45,7 ± 1,79	88,6 ± 1,19
	Random Forest	43,3 ± 1,76	92,0 ± 1,02
CONTROL	kNN	50,5 ± 1,47	74,1 ± 1,30
	SVM	57,6 ± 1,45	92,1 ± 0,82
	Red neuronal	59,1 ± 1,45	91,3 ± 0,86
	Random Forest	54,7 ± 1,46	93,3 ± 0,76
EVOLUCION	kNN	45,7 ± 1,82	89,1 ± 1,19
	SVM	49,7 ± 1,83	95,4 ± 0,82
	Red neuronal	57,4 ± 1,82	95,4 ± 0,82
	Random Forest	64,0 ± 1,82	94,3 ± 0,90

Cuadro 4: Rendimiento de clasificación en identificación y autenticación para las palabras QUE, DESDE, CONTROL y EVOLUCION.

### 4.3. Grupos naturales de tipistas

Las cadencias de tecleo pueden agruparse naturalmente de diversas maneras. Una división evidente deriva de la mano dominante del usuario, zurdo o diestro. También se puede postular la existencia de estratos, o quizás un espectro, de competencia que abarca desde el que escribe con un único dedo hasta los expertos mecanógrafos. Es más sutil la distinción que puede hacer en base al estado emocional del individuo en el momento de la escritura [5,8,2]. A los fines de evaluar qué tipos de grupos naturales pueden extraerse del conjunto de datos dado, se evaluó la aplicación de mapas auto-organizados SOM, una técnica de agrupamiento y aprendizaje no supervisado. La aplicación de mapas auto-organizados sobre los 158 usuarios del dataset revela un fenómeno inesperado: en la Figura 3 se observan los resultados de dos ejecuciones con una grilla de  $5 \times 5$  e inicialización aleatoria no replicable. Nótese que en ambos casos aparecen dos grupos destacados con casi idéntica cantidad de instancias, seguidos de lejos por diversos grupos menores, entre los cuáles uno sólo parece destacar secundariamente. Los resultados no varían significativamente si se utilizan otros tamaños de grilla u otros tipos de inicialización. Si bien no se cuenta con una clasificación confiable de la mano dominante del usuario en el dataset desestructurado de origen desde el cuál se han obtenido los datos tabulares que aquí se estudian, y que permitiría confirmar esta hipótesis, el tamaño del grupo secundario dominante es consistente con la cantidad de zurdos en la población general, general el orden del 10%.



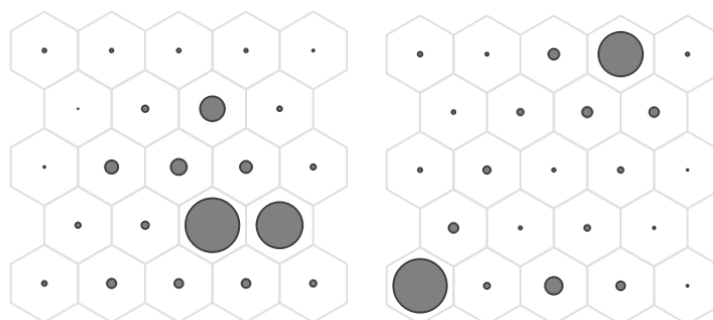


Figura 3: Dos agrupamientos con SOM e inicialización aleatoria

## 5. Conclusiones

En este trabajo se han explorado las correlaciones internas de parámetros temporales, el rendimiento de clasificación en las tareas de identificación y verificación, y se buscaron grupos naturales de tipistas, entre los que se encuentran diestros y zurdos, en un subconjunto del dataset de [12], que contiene multitud de sesiones de escritura en un teclado convencional donde se han registrado las teclas presionadas y los tiempos de retención y latencia. Se observó que la influencia de cada parámetro temporal no es necesariamente local e inmediata, limitada a la tecla siguiente o una cercana, sino que pueden encontrarse correlaciones maximales entre letras con cuatro y hasta cinco grados de separación como la E y la C, o la  $O_1$  y la  $O_2$  en la palabra EVOLUCION. La hipótesis de existencia de correlaciones fuertes entre teclas no adyacentes, planteada en la Sección 2 se mostró particularizable también a los usuarios individuales en el Cuadro 3. El rendimiento en la clasificación, tanto en tareas de identificación como de verificación, resultó sorprendentemente bueno a pesar de la simplicidad del esquema utilizado, particularmente en las palabras más largas —EVOLUCION y CONTROL—, pero no debe olvidarse que las palabras elegidas son las más comunes del dataset para el largo dado y palabras más raras no exhibirían métricas de precisión tan optimistas. Sin embargo, estos resultados son lo suficientemente alentadores como para proponer la posibilidad de identificar el creador de un texto libre fragmentando la sesión por palabras y utilizando clasificadores individuales para cada una de ellas. Finalmente, la aplicación de mapas auto-organizados reveló un fenómeno inesperado; dos grupos destacados con casi idéntica cantidad de instancias, que no varían significativamente si se utilizan otros tamaños de grilla u otros tipos de inicialización. En el contexto de las futuras líneas de trabajo se contempla profundizar las pruebas de identificación de usuarios, analizar el empleo de mapas auto organizados para la terminación de poblaciones de tipo secundarias de grupos naturales de tipistas como medida para fortalecer la identificación y autenticación de usuarios.

## Referencias

1. Bergadano, F., Gunetti, D., Picardi, C.: User authentication through keystroke dynamics. *ACM Transactions on Information and System Security (TISSEC)* 5(4), 367–397 (2002)
2. Calot, E.P.: Robustez de las métricas de clasificación de cadencia de tecleo frente a variaciones emocionales. Ph.D. thesis, Universidad Nacional de La Plata (2019), <http://sedici.unlp.edu.ar/handle/10915/76652>
3. Calot, E.P., Ierache, J.S.: Multimodal biometric recording architecture for the exploitation of applications in the context of affective computing. In: *Proceedings del XXIII CACIC (La Plata, 2017)* (2017)
4. Calot, E.P., Ierache, J.S., Hasperué, W.: Robustness of keystroke dynamics identification algorithms against brain-wave variations associated with emotional variations. In: *Advances in Intelligent Systems and Computing*. Springer (2019)
5. Calot, E.P., Rodríguez, J.M., Pereira, G., Iribarren, J., Ierache, J.S.: Métodos adaptativos de educación de dinámica de tecleo centrado en el contexto emocional de un individuo aplicando interfaz cerebro computadora. In: *XVI WICC* (2014)
6. Concilio, G.M.: Aplicación de técnicas de modelado de la cadencia de tecleo para fortalecer la identificación del usuario en el comercio electrónico (Feb 2019)
7. Concilio, G.M., Calot, E.P., Ierache, J.S., Merlino, H.D.: Application of Keystroke Dynamics Modelling Techniques to Strengthen the User Identification in the Context of E-commerce (2018)
8. Epp, C., Lippold, M., Mandryk, R.L.: Identifying emotional states using keystroke dynamics. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 715–724. ACM (2011)
9. Forsen, G.E., Nelson, M.R., Staron, R.J.J.: Personal attributes authentication techniques. Tech. rep., DTIC Document (October 1977), <http://lsia.fi.uba.ar/papers/forsen77.pdf>, final technical rept.
10. Francesco Bergadano, D.G.y.C.P.: Identity verification through dynamic keystroke analysis. *Intelligent Data Analysis* 7(5), 469–496 (2003)
11. González, N.: Finite context modelling of keystroke dynamics in mixed authentication schemes (2014), <http://lsia.fi.uba.ar/papers/gonzalez14b.pdf>, LSIA, FIUBA, Argentina
12. González, N., Calot, E.P., Ierache, J.S.: A replication of two free text keystroke dynamics experiments under harsher conditions. In: *Biometrics Special Interest Group (BIOSIG), 2016 International Conference of the*. pp. 1–6. IEEE (2016)
13. González, N., Concilio, G., Ierache, J., Hasperué, W.: <https://drive.google.com/drive/folders/1T1MUcCivvMrmBTxwTHzpKdW8zZQ3INK?usp=sharing>
14. Idrus, S.Z.S., Cherrier, E., Rosenberger, C., Bours, P.: Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords. *Computers & Security* 45, 147–155 (2014)
15. Monroe, F., Reiter, M.K., Wetzel, S.: Password hardening based on keystroke dynamics. In: *Proceedings of the 6th ACM Conference on Computer and Communications Security*. pp. 73–82. CCS '99, ACM, New York, NY, USA (1999), <http://doi.acm.org/10.1145/319709.319720>
16. Morales, A., Acien, A., Fierrez, J., Monaco, J.V., Tolosana, R., Vera-Rodríguez, R., Ortega-García, J.: Keystroke biometrics in response to fake news propagation in a global pandemic. *arXiv preprint arXiv:2005.07688* (2020)
17. Van Waes, L., Leijten, M., Mariën, P., Engelborghs, S.: Typing competencies in alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior* 73, 311–319 (2017)