


# Atributos derivados para la clasificación de cadencias de tecleo en textos libres basados en el grado de desorden

Nahuel González

 [orcid.org/0000-0001-5570-6922](https://orcid.org/0000-0001-5570-6922)

Laboratorio de Sistemas de Información Avanzados (LSIA),  
Facultad de Ingeniería, Universidad de Buenos Aires,  
Ciudad Autónoma de Buenos Aires, Argentina  
[ngonzalez@lsia.fi.uba.ar](mailto:ngonzalez@lsia.fi.uba.ar)

**Resumen** Se proponen cuatro atributos derivados (*visibilidad, posición, y desorden*) para comparar muestras de cadencias de tecleo en textos libres, basados en el grado de desorden de los tiempos de retención y latencia, y se evalúa su rendimiento. Al verificar la identidad de los usuarios resultan en un FAR del 3,1% y un FRR del 4,6%, similares al de las métricas del estado del arte, y sin problemas de colinealidad. Al complementar a las anteriores, reducen el FAR y el FRR hasta el 2%.

**Keywords:** seguridad informática, biometría comportamental, cadencias de tecleo, texto libre, aprendizaje automático, atributos derivados

## 1. Introducción

El análisis de los patrones característicos de escritura en un teclado convencional o un dispositivo móvil configura un dominio de la biometría comportamental denominado *cadencias de tecleo*. Dentro del ámbito de la seguridad informática, las cadencias de tecleo han sido ante todo empleadas como un segundo factor de autenticación (2/3FA) para la verificación de la identidad de los usuarios [1], aunque también han sido utilizadas, luego de ser extraídas en un ataque por canal lateral que revela la cadencia de tecleo pero no el texto ingresado, para reconstruir este último en base a los tiempos de escritura [2, 3] y para reducir la complejidad de un ataque por fuerza bruta [4].

Aunque el estado del arte en la verificación de usuarios por medio de cadencias de tecleo utiliza técnicas de aprendizaje profundo [5], recientemente se ha demostrado que puede mejorarse el rendimiento de clasificación en esta y otras modalidades de biometría comportamental, a la vez que se reduce el tiempo de entrenamiento y el tamaño del conjunto de datos necesario, si se complementa a las redes neuronales profundas con atributos derivados propios del dominio [6]. Más específicamente, estos han resultado efectivos para la detección de vida, con el objetivo de proteger a los sistemas de autenticación basados en cadencias de tecleo contra ataques de presentación con muestras sintetizadas [7].

En particular, la métrica R [8] ha demostrado tener un rendimiento excelente en diversos estudios comparativos [9, 10], aunque tiene la desventaja de requerir muestras grandes, del orden de 700 teclas, para alcanzar el rendimiento óptimo. La misma se basa en el grado de desorden global de los tiempos de escritura; ordena los tiempos entre teclas de las muestras a comparar y mide cuánto varían las posiciones relativas. Este idea no ha sido generalizada en la literatura del tema y, específicamente, no se ha estudiado cómo reducir la cantidad de teclas necesarias, o como aplicarla localmente a fragmentos de la muestra.

**Contribuciones.** El objeto de este estudio es proponer y evaluar el rendimiento de cuatro nuevos atributos basados en el grado de desorden, especializados para la verificación de cadencias de tecleo en textos libres. Las principales contribuciones ofrecidas son:

- Se proponen cuatro nuevos atributos basadas en el grado de desorden de los tiempos de escritura para comparar muestras de cadencias de tecleo.
- Se evaluó el rendimiento de los atributos propuestos para la verificación de identidad sobre un conjunto de datos realista, públicamente accesible, y que ha sido utilizado en estudios previos del tema [11].
- Se ofrece en forma abierta y gratuita el conjunto de datos de entrenamiento [11] para facilitar la verificación de los resultados.

**Organización.** El resto del artículo está organizado como se describe a continuación. La sección 2 reseña algunos estudios previos sobre el tema. La sección 3 describe los métodos propuestos. La sección 4 detalla la metodología del experimento. La sección 5 discute los resultados. La sección 6 resume las conclusiones.

## 2. Estudios previos

La métrica R de Bergadano, Gunetti, y Picardi [1] ha representado un hito en la disciplina, permitiendo por primera vez la verificación de cadencias de tecleo en el análisis de textos libres con bajas tasas de error. En el mismo artículo han propuesto también la métrica A, cuyo rendimiento no generaliza favorablemente al ser evaluada bajo otro conjunto de datos [12]. Killourhy y Maxion [9] comparan una decena de métricas y atributos, entre ellos  $\mathcal{L}_1$  y  $\mathcal{L}_2$ , que se utilizan como parte del control en este experimento. El conteo de valores atípicos, denominado Z y también utilizado como control, fue propuesto por Haider y colaboradores [13]; aunque su rendimiento individual es pobre, suele ser útil para complementar a otros atributos en un esquema de clasificación debido a la baja correlación con ellos. Hasta donde alcanza nuestro conocimiento de la literatura del tema no se han propuesto, con la excepción de la métrica R, otros atributos basados en el grado de desorden global de los tiempos de escritura.

## 3. Métodos propuestos

### 3.1. Definiciones

Una muestra  $M$  de largo  $l = l(M) = l(M')$  contiene una secuencia de teclas  $k_1 \dots k_l$  y sus correspondientes tiempos  $t_1 \dots t_l$ , de retención (intervalo entre

eventos de presión y liberación) o latencia (intervalo entre presión de teclas sucesivas). Denominamos  $k_i$  a la  $i$ -ésima tecla de la muestra  $M$ ; su *contexto de orden  $m$*  es la secuencia  $k_{i-m} \dots k_i$  de teclas que la anteceden, junto con  $k_i$ . Al comparar dos muestras  $M$  y  $M'$  supondremos siempre que el largo y las secuencias de teclas de ambas son idénticas; cuando sea necesario, utilizaremos  $t'_1 \dots t'_l$  para referirnos a los tiempos de la muestra  $M'$ .

### 3.2. V (visibilidad)

Diremos que la tecla  $k_i$  tiene *visibilidad descendente hasta  $n$  teclas atrás en la muestra  $M$* , que denotamos con  $v^-(M, i) = n$ , si se cumplen simultáneamente las dos condiciones

$$(a) t_i \leq t_{i-n} \quad (b) t_i > t_{i-j} \quad \forall 0 < j < n$$

Idénticamente, diremos que la tecla  $k_i$  tiene *visibilidad ascendente hasta  $n$  teclas atrás en la muestra  $M$* , que denotamos con  $v^+(M, i) = n$ , si se cumple simultáneamente que

$$(a) t_i \geq t_{i-n} \quad (b) t_i < t_{i-j} \quad \forall 0 < j < n$$

Dada una cota  $N$ , el *atributo  $V^-$  de visibilidad descendente* cuantifica la diferencia entre las muestras  $M$  y  $M'$ , ambas de largo  $l$ , en la siguiente forma

$$V^-(M, M') = \frac{1}{N(l-N)} \sum_{i=N+1}^l |\min\{N, v^-(M, i)\} - \min\{N, v^-(M', i)\}| \quad (1)$$

que devuelve un valor entre 0 (cuando ambas muestras son idénticas) y 1.

Utilizando  $v^+$  en la expresión de arriba en lugar de  $v^-$  obtenemos la expresión para  $V^+$ , el *atributo de visibilidad ascendente*. El objetivo de la cota  $N$  y los máximos dentro de la expresión es restringir la influencia de cada tecla al ámbito local y conservarla independiente del tamaño de la muestra; ya que de lo contrario, las teclas con mínimo y máximo tiempo en la muestra pueden tener una contribución exagerada si se encuentran cerca del final de la misma.

### 3.3. P (posición)

La *posición* de una tecla  $k_i$  en su contexto de orden  $m$ , que denotamos con  $p(M, m, i)$ , es el índice que le corresponde luego de ordenar las teclas del contexto en base a sus tiempos. Formalmente, el ordenar ambas filas de

$$\begin{array}{ccccccc} k_{i-m} & k_{i-m+1} & \dots & k_{i-1} & k_i \\ t_{i-m} & t_{i-m+1} & \dots & t_{i-1} & t_i \end{array}$$

utilizando los valores temporales de  $t_{i-m} \dots t_i$  resulta en una biyección  $\sigma : [0, m] \rightarrow [i-m, i]$  tal que  $t_{\sigma(0)} \leq t_{\sigma(1)} \leq \dots \leq t_{\sigma(m-1)} \leq t_{\sigma(m)}$  y

$$\begin{array}{ccccccc} k_{\sigma(0)} & k_{\sigma(1)} & \dots & k_{\sigma(m-1)} & k_{\sigma(m)} \\ t_{\sigma(0)} & t_{\sigma(1)} & \dots & t_{\sigma(m-1)} & t_{\sigma(m)} \end{array}$$

Utilizando esta biyección, definimos

$$p(M, m, i) = \sigma^{-1}(i)$$

Por ejemplo, los tiempos de latencia al escribir HOLA, que es el contexto de orden 3 de la tecla A, pueden haber sido

$$\begin{array}{cccc} H & O & L & A \\ 153 & 278 & 176 & 190 \end{array}$$

Luego de ordenar por la segunda fila obtenemos

$$\begin{array}{cccc} H & L & A & O \\ 153 & 176 & 190 & 278 \end{array}$$

Y la posición de la tecla A en su contexto de orden 3, HOLA, es igual a 2 (contando desde cero) porque su valor temporal es mayor que el de la H y el de la L, pero menor que el de la O. Ahora, fijamos el orden del contexto a un determinado valor  $N$  y cuantificamos la diferencia entre las muestras  $M$  y  $M'$  utilizando el atributo  $P$  de posición en la siguiente forma

$$P(M, M') = \frac{1}{N(l-N)} \sum_{i=N+1}^l |p(M, N, i) - p(M', N, i)| \quad (2)$$

que devuelve un valor entre 0 (cuando ambas muestras son idénticas) y 1.

### 3.4. D (desorden)

Una vez más, consideremos la biyección  $\sigma_i : [0, m] \rightarrow [i-m, i]$  que resulta de ordenar las teclas del contexto de orden  $m$  de la tecla  $k_i$  en base a los valores de tiempo  $t_{i-m} \dots t_i$  en la muestra  $M$ .

$$\begin{array}{cccc} k_{\sigma(0)} & k_{\sigma(1)} & \dots & k_{\sigma(m-1)} & k_{\sigma(m)} \\ t_{\sigma(0)} & t_{\sigma(1)} & \dots & t_{\sigma(m-1)} & t_{\sigma(m)} \end{array}$$

El mismo proceso aplicado a la muestra  $M'$  resulta en otra biyección, llamémosla  $\tau_i$ , que puede o no ser igual a  $\sigma_i$  ya que  $M$  y  $M'$  tienen igual secuencia de teclas pero no necesariamente de tiempos. Así, luego de ordenar, obtenemos

$$\begin{array}{cccc} k_{\tau(0)} & k_{\tau(1)} & \dots & k_{\tau(m-1)} & k_{\tau(m)} \\ t'_{\tau(0)} & t'_{\tau(1)} & \dots & t'_{\tau(m-1)} & t'_{\tau(m)} \end{array}$$

Ahora queremos medir el *grado de desorden local*, es decir cuánto difieren entre sí las biyecciones  $\sigma$  y  $\tau$  para un contexto particular. El problema ha sido estudiado por Diaconis y Graham [14], que proponen diversas métricas de desorden. Entre otras, consideran

$$T(\sigma, \tau) = \text{mínimo de transposiciones necesarias en la permutación que transforma } (\sigma(0), \dots, \sigma(m)) \text{ en } (\tau(0), \dots, \tau(m))$$

Los mismos autores proveen un algoritmo de cálculo rápido basado en el número de ciclos en la permutación  $\sigma^{-1}\tau$ , denominado  $C(\sigma^{-1}\tau)$ , obteniendo

$$T(\sigma, \tau) = m + 1 - C(\sigma^{-1}\tau)$$

Fijamos el orden del contexto a un determinado valor  $N$ ; el máximo de transposiciones posibles en una permutación de orden  $N + 1$  es  $N$ . Entonces, en base a la expresión anterior podemos definir el *atributo D de desorden local* para cuantificar la diferencia entre las muestras  $M$  y  $M'$ , en la forma

$$D(M, M') = \left\lfloor \frac{N}{l} \right\rfloor \sum_{i=1}^{\lfloor l/N \rfloor} \frac{T(\sigma_{iN}, \tau_{iN})}{N} \quad (3)$$

donde  $\lfloor \cdot \rfloor$  denota la función piso,  $\sigma_i$  es la biyección que corresponde a ordenar los tiempos del contexto de orden  $m$  de  $k_i$  en la muestra  $M$ , y  $\tau_i$  en la muestra  $M'$ . Al igual que las anteriores, devuelve un valor entre 0 (para muestras que son idénticas) y 1.

## 4. Evaluación experimental

### 4.1. El conjunto de datos

La evaluación de los métodos propuestos se realizó con el conjunto de datos LSIA [15, 12], que contiene muestras de cadencias de tecleo en texto libre. Estas incluyen las secuencias de códigos de teclas presionadas, tiempos de retención (intervalo entre presión y liberación de tecla), y latencias (intervalo entre presión de teclas sucesivas). El conjunto de datos cuenta con 136 usuarios y 13.600 sesiones, de entre 150 y más de 1000 caracteres cada una. Se encuentra a disponibilidad del público en forma abierta y gratuita [11].

### 4.2. Preprocesamiento y limpieza de los datos

Se utilizó una herramienta propietaria desarrollada por el LSIA para preprocesar las muestras del conjunto de datos, con el objetivo de eliminar valores inválidos o demasiado grandes que correspondan a pausas externas y no al ritmo natural de escritura. En particular, todos los valores mayores a 1500mseg. fueron marcados como inválidos, con el objetivo de que no sean incluidos en el cómputo de los atributos a evaluar.

### 4.3. Optimización de parámetros

Los atributos V, P, y D, descritos en la sección 3.4, requieren fijar un valor para el parámetro  $N$ , que determina el orden de los contextos considerados. No es evidente *a priori* qué valor resultaría óptimo, ameritando una búsqueda exhaustiva. Con el objetivo de optimizar el valor de  $N$  se llevó a cabo el proceso de evaluación del EER que se describirá en la siguiente sección, sobre la totalidad del conjunto de datos, pero con distintos valores de  $N$ . En particular, se evaluó  $2 \leq N \leq 10$  para los atributos  $V^+$  y  $V^-$ ,  $3 \leq N \leq 10$  para  $P$ , y  $4 \leq N \leq 20$  para  $D$ .

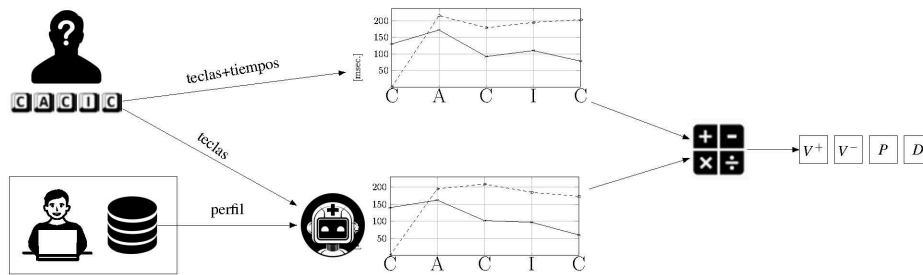


Figura 1: Esquema de evaluación de los atributos  $V^+$ ,  $V^-$ ,  $P$ , y  $D$  entre una muestra dada y su reconstrucción por contextos finitos

#### 4.4. Evaluación del rendimiento individual de los atributos

En la sección 3.4 se ha indicado que los atributos  $V^+$ ,  $V^-$ ,  $P$ , y  $D$  sólo pueden comparar muestras con idéntica secuencia de teclas. Dado una muestra de un usuario cuya identidad se desea verificar, al tratar con texto libre es poco probable que el perfil del usuario cuente con una muestra similar. Por este motivo es necesario reconstruir cómo el usuario legítimo habría escrito el texto en cuestión. Para tal fin se utilizará el método de *modelado por contextos finitos* [15]. Este método, dada una secuencia de teclas, reconstruye la cadencia de tecleo del usuario legítimo; utiliza las muestras existentes en su perfil buscando para cada tecla observaciones anteriores de sus tiempos de retención y latencia, y prefiriendo aquellas observaciones cuyas teclas precedentes, o *contexto*, se corresponden con las de la muestra a reconstruir. Este método ha sido empleado previamente para la autenticación de usuarios por medio de su cadencia de tecleo [15], para la generación de muestras sintéticas y la detección de vida [7], y para la reconstrucción de un texto en base a los tiempos de escritura [3]. Para una descripción detallada se recomienda consultar las referencias.

La figura 1 muestra el esquema de evaluación de los atributos  $V^+$ ,  $V^-$ ,  $P$ , y  $D$ . Dada una muestra  $M$  de un usuario desconocido, se construye otra muestra  $M'$  con la misma secuencia de teclas, en base al perfil del usuario legítimo y utilizando contextos finitos. Como ambas muestras comparten la secuencia de teclas, pueden utilizarse las ecuaciones (1), (2), y (3) para calcular  $V$ ,  $D$ , y  $P$ .

Cada usuario cuenta con 100 muestras en el conjunto de datos; de estas, 50 al azar se utilizaron como perfil para el modelado con contextos finitos. Las restantes 50 muestras de cada usuario legítimo, junto con 50 muestras adicionales de impostores tomadas al azar entre todas las muestras del resto de los usuarios, fueron evaluadas según el esquema anterior. Los 100 valores de  $V$ ,  $D$ , y  $P$  resultantes fueron utilizados para trazar las curvas FAR/FRR y obtener el EER.

#### 4.5. Evaluación del rendimiento conjunto

La utilización conjunta de diversos atributos derivados disminuye el error de clasificación si estos provienen suficiente ganancia de información y no se encuen-

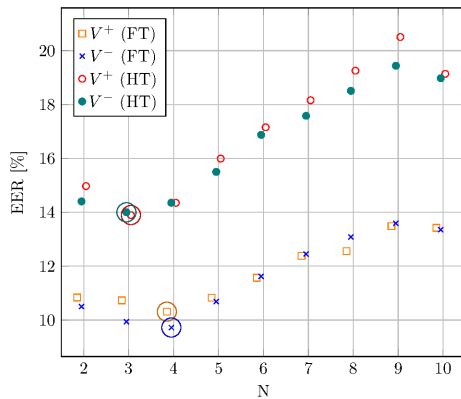


Figura 2: Rendimiento de  $V^+$  y  $V^-$  con distintos valores de  $N$ .

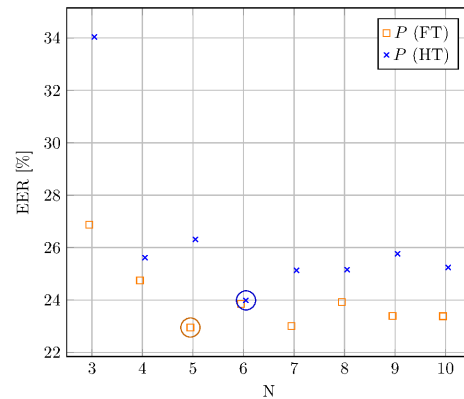


Figura 3: Rendimiento de  $P$  con distintos valores de  $N$ .

tran fuertemente correlacionados. Para evaluar el rendimiento conjunto de los atributos propuestos, se entrenó un clasificador SVM binario para verificar la identidad de los usuarios, con los parámetros que se indican en la Sección 4.6.

Se utilizaron los valores de los cuatro atributos propuestos ( $V^+$ ,  $V^-$ ,  $P$ ,  $D$ ), calculados tanto para tiempos de retención como para latencias, totalizando ocho atributos. El conjunto de entrenamiento se compuso de los valores correspondientes a las 50 muestras del perfil del usuario (ver sección anterior), marcadas como legítimas; además de 50 muestras adicionales tomadas al azar entre el resto de los usuarios, marcadas como impostores.

A modo de experimento comparativo, se evaluó idénticamente el rendimiento conjunto de cuatro métricas bien establecidas y de reconocido rendimiento:  $\mathcal{L}_1$ ,  $\mathcal{L}_2$ ,  $R$ , y  $Z$ , también para tiempos de retención y latencias. Finalmente, se evaluó el rendimiento conjunto de las métricas propuestas y de control, además de determinar la correlación entre todas ellas, para determinar si las primeras son útiles para reducir el error de clasificación.

#### 4.6. Materiales y herramientas

Como se ha descrito en la Sección 4.2, se utilizó una herramienta propietaria desarrollada por el LSIA para preprocesar y limpiar las muestras del conjunto de datos. Para el modelado de cadencias de tecleo por medio de contextos se utilizó la implementación disponible en [16]. La clasificación se realizó con la implementación SVM de WEKA 3.8.4 [17], utilizando optimización secuencial mínima (SMO) [18] para el entrenamiento, núcleo polinomial, y calibrador logístico con parámetros por defecto.

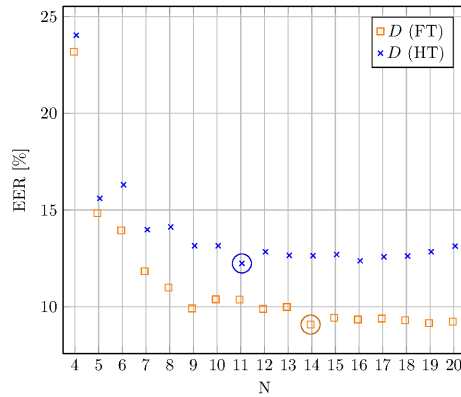


Figura 4: Rendimiento de  $D$  con distintos valores de  $N$ .

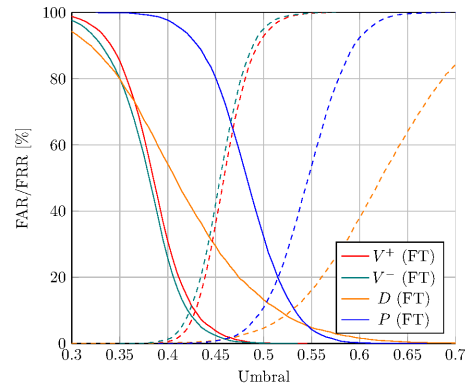


Figura 5: Curvas de FAR/FRR para latencias. Impostores en línea punteada.

## 5. Resultados y discusión

Los resultados de la optimización del parámetro  $N$  para los atributos  $V^+$ ,  $V^-$ ,  $P$ , y  $D$  se muestran en las figuras 2, 3, y 4, respectivamente. Los mínimos de EER para  $V^+$  y  $V^-$  se obtienen con  $N = 3$  para tiempos de retención y  $N = 4$  para latencias. En el caso de  $P$ , los valores óptimos son  $N = 6$  para tiempos de retención y  $N = 5$  para latencias, mientras que para  $D$  son  $N = 11$  y  $N = 14$ , respectivamente. Un detalle de las curvas de FAR/FRR para latencia con estos valores de  $N$  se muestran en la figura 5.

Conjunto	Atributos	FAR	FRR
Propuestas	$V^+$ , $V^-$ , $P$ , $D$	3,1% ( $\pm 0,7$ )	4,6 ( $\pm 1,5$ )
Control	$\mathcal{L}_1$ , $\mathcal{L}_2$ , $R$ , $Z$	3,9% ( $\pm 0,7$ )	3,9% ( $\pm 0,4$ )
Todos	$V^+$ , $V^-$ , $P$ , $D$ , $\mathcal{L}_1$ , $\mathcal{L}_2$ , $R$ , $Z$	<b>2,2% (<math>\pm 0,5</math>)</b>	<b>1,9% (<math>\pm 0,6</math>)</b>

Cuadro 1: Rendimiento conjunto de los atributos propuestos y las métricas de control

Con los valores óptimos de  $N$ , el rendimiento conjunto en la clasificación de los atributos propuestos se muestra en la tabla 1, junto con el rendimiento bajo el mismo esquema de las métricas de control y de ambos tipos simultáneamente, todas con intervalos de confianza del 95%. Los valores de FAR y FRR obtenidos con los atributos propuestas muestran que las mismas son competitivas con las del estado del arte, ambas en el rango del 3–5%; como los intervalos de confianza se solapan, la diferencia no tiene significación estadística. Lo que es más, la utilización conjunta reduce el error de clasificación hasta cerca del 2%, aproxi-



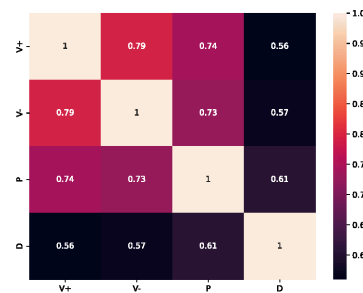
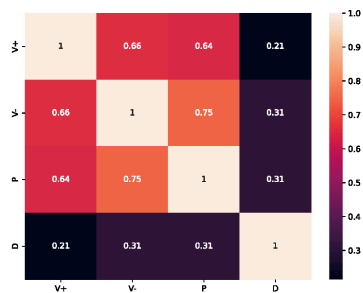


Figura 6: Matriz de correlación de  $V^+$ ,  $V^-$ ,  $P$ , y  $D$  para latencias. Figura 7: Matriz de correlación de  $V^+$ ,  $V^-$ ,  $P$ , y  $D$  para tiempos de retención.

madamente la mitad de los casos anteriores, y aquí los intervalos de confianza no se solapan por lo que sí hay significación estadística.

Las figuras 6 y 7 muestran las matrices de correlación de valores de los atributos propuestos para tiempos de retención y de latencia. En general, las correlaciones son mayores para tiempos de retención que para latencias. Como es esperable por su similitud, los atributos  $V^+$  y  $V^-$  son las que muestran mayor correlación entre sí, pero ningún par de ellas supera el valor de  $r = 0,79$ . Un estudio reciente sobre los efectos negativos de la colinealidad en modelos de aprendizaje automático [19] establece un valor de corte aceptable de  $r^2 = 0,6$  para evitar problemas, que corresponde a  $r \approx 0,78$ . Todos los pares de atributos, con la excepción de  $V^+$  y  $V^-$  para tiempos de retención que se encuentra en el límite, presentan valores de correlación inferiores.

**Limitaciones de este estudio.** El conjunto de datos LSIA incluye sólo muestras en idioma castellano. Generalizar el rendimiento de los atributos propuestos a otros idiomas requerirá conjuntos de datos adicionales.

## 6. Conclusión

En el presente artículo se han propuesto cuatro atributos para comparar muestras de cadencias de tecleo en textos libre, basadas en el grado de desorden de los tiempos de retención y latencia, que exploran la idea subyacente en la métrica R [8] intentando aplicarla localmente y a muestras más reducidas. Individualmente, el atributo  $D$  para latencias con  $N = 14$  alcanza el EER más bajo, de aproximadamente 10%. Sobre el conjunto de datos utilizado para la evaluación, el empleo de los atributos propuestos para la verificación de la identidad de los usuarios resulta en un FAR del 3,1% y un FRR del 4,6%. En contraste, las métricas del estado del arte alcanzan un FAR y un FRR del 3,9%, que luego de considerar intervalos de confianza resultan similares a los anteriores. Los atributos propuestos mostraron correlaciones mutuas suficientemente bajas

como para evitar los problemas acarreados por la colinealidad, y al complementar a las métricas del estado del arte redujeron el FAR y el FRR hasta casi el 2%.

**Futuras líneas de investigación.** Los atributos propuestos prometen ser aplicables para la detección de vida y la discriminación de muestras sintetizadas. La exploración de estos usos queda relegado a futuras líneas de investigación.

## Bibliografía

- [1] Francesco Bergadano, Daniele Gunetti, and Claudia Picardi. User authentication through keystroke dynamics. *ACM Transactions on Information and System Security (TISSEC)*, 5(4): 367–397, 2002.
- [2] John V Monaco. What are you searching for? a remote keylogging attack on search engine autocomplete. In *28th {USENIX} Security Symposium ({USENIX} Security 19)*, pages 959–976, 2019.
- [3] Nahuel González, Enrique P. Calot, Jorge S. Ierache, and Waldo Hasperux00E9;. The reverse problem of keystroke dynamics: Guessing typed text with keystroke timings only. In *2021 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, pages 1–6, 2021. doi:10.1109/ICECET52533.2021.9698782.
- [4] 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.
- [5] Alejandro Acien, Aythami Morales, Ruben Vera-Rodriguez, Julian Fierrez, and John V Monaco. Typenet: Scaling up keystroke biometrics. In *2020 IEEE International Joint Conference on Biometrics (IJC/B)*, pages 1–7. IEEE, 2020.
- [6] Sakorn Mekruksavanich and Anuchit Jitpattanakul. Biometric user identification based on human activity recognition using wearable sensors: An experiment using deep learning models. *Electronics*, 10(3):308, 2021.
- [7] Nahuel González, Enrique P. Calot, Jorge S. Ierache, and Waldo Hasperué. Towards liveness detection in keystroke dynamics: Revealing synthetic forgeries. *Systems and Soft Computing*, 4:200037, 2022. ISSN 2772-9419. doi:https://doi.org/10.1016/j.sasc.2022.200037. URL https://www.sciencedirect.com/science/article/pii/S2772941922000047.
- [8] Daniele Gunetti and Claudia Picardi. Keystroke analysis of free text. *ACM Transactions on Information and System Security (TISSEC)*, 8(3):312–347, 2005.
- [9] Kevin S Killourhy and Roy A Maxion. Comparing anomaly-detection algorithms for keystroke dynamics. In *2009 IEEE/IFIP International Conference on Dependable Systems & Networks*, pages 125–134. IEEE, 2009.
- [10] Jiaju Huang, Daqing Hou, Stephanie Schuckers, Timothy Law, and Adam Sherwin. Benchmarking keystroke authentication algorithms. In *2017 IEEE Workshop on Information Forensics and Security (WIFS)*, pages 1–6. IEEE, 2017.
- [11] Enrique P. Calot. Keystroke dynamics keypress latency dataset. Database, jan 2015. URL <http://lsia.fi.uba.ar/pub/papers/kd-dataset/>.
- [12] Nahuel González, Enrique P Calot, and Jorge S Ierache. A replication of two free text keystroke dynamics experiments under harsher conditions. In *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–6. IEEE, 2016.
- [13] Ahmed Abbas y Abbas K. Zaidi Sajjad Haider. A multi-technique approach for user identification through keystroke dynamics. In *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, volume 2, pages 1336–1341. IEEE, 2000.
- [14] Persi Diaconis and Ronald L Graham. Spearman’s footrule as a measure of disarray. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(2):262–268, 1977.
- [15] Nahuel González and Enrique P Calot. Finite context modeling of keystroke dynamics in free text. In *Biometrics Special Interest Group (BIOSIG), 2015 International Conference of the*, pages 1–5. IEEE, 2015.
- [16] Nahuel González. Herramienta de análisis de cadencias de tecleo en texto libre, 7 2021. URL <https://github.com/lsia/herramientaGonzalez2021>.
- [17] Ian H Witten, Eibe Frank, Mark A Hall, CJ Pal, and MINING DATA. Practical machine learning tools and techniques. In *DATA MINING*, volume 2, page 4, 2005.
- [18] John Platt. Sequential minimal optimization: A fast algorithm for training support vector machines. 1998.
- [19] Ron Johnston, Kelvyn Jones, and David Manley. Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of british voting behaviour. *Quality & quantity*, 52(4):1957–1976, 2018.