
MÁSTER EN INTELIGENCIA ARTIFICIAL

# DESARROLLO DE UN SISTEMA DE AUTENTICACIÓN BASADO EN LA BIOMETRÍA DEL COMPORTAMIENTO EN DISPOSITIVOS MÓVILES

*Curso 2025/26 -*

**Trabajo dirigido por:**
ROBERTO ALCARAZ MACHADO

**DAVID JIMÉNEZ CASTRO**

**DNI:** 76052663-N
**e-mail:** djimenezc1@student.universidadviu.com

# Índice general

# Capítulo 1

# Introducción

En el panorama actual de la ciberseguridad móvil, la dependencia de métodos de autenticación estática —como contraseñas, PINs o patrones de dibujo— presenta vulnerabilidades críticas frente a ataques de ingeniería social e interceptación. Si bien la biometría fisiológica (huella dactilar o reconocimiento facial) ha reforzado el perímetro de seguridad, estos métodos actúan como guardianes de "un solo paso" que no protegen la sesión una vez iniciada.

El presente proyecto propone el desarrollo de un **Sistema de Autenticación Basado en la Biometría del Comportamiento**, un enfoque que permite la verificación de identidad de forma **pasiva, continua y transparente** para el usuario.

### 1.0.1. El Fundamento de la Biometría Conductual

A diferencia de los rasgos físicos, la biometría conductual se centra en **cómo** interactúa el usuario con su dispositivo[cite: 26]. Este sistema aprovecha la singularidad de los patrones neurofisiológicos y motores, analizando variables clave capturadas de manera transparente:

- **Dinámica de Tecleo (*Keystroke Dynamics*):** Estudio de los tiempos de presión (*Dwell Time*), las latencias entre teclas (*Flight Time*) y los intervalos entre pulsaciones (*Inter-key Interval*).
- **Interacción Táctil:** Patrones de desplazamiento (*scrolling*), toques en pantalla (*tapping*) y gestos en la interfaz.
- **Sensores de Movimiento (*IMU*):** Uso de acelerómetros, giroscopios y magnetómetros para capturar la micro-gestualidad y el manejo físico del terminal durante la interacción.

### 1.0.2. Innovación Tecnológica y Arquitectura

Para superar las limitaciones de los modelos clásicos (como SVM o KNN), que presentan dificultades de escalabilidad en entornos masivos [cite: 86, 87], este sistema se fundamenta en arquitecturas de **Aprendizaje Profundo (*Deep Learning*)**:

- **Modelos Híbridos y Atención:** Integración de redes neuronales convolucionales (CNN) para la extracción de características locales y redes recurrentes (LSTM) para dependencias temporales globales.
- **Mecanismos de Atención Dual:** Empleo de arquitecturas basadas en *transformers* (como *STDAT*) que aplican atención tanto temporal como de canal para capturar patrones característicos.
- **Fusión Multimodal:** Combinación de la dinámica de tecleo con datos de sensores inerciales para incrementar la robustez del sistema frente a cambios de contexto.
- **Funciones de Pérdida Avanzadas:** Implementación de técnicas como *Triplet Loss* y *Set2set Loss*, permitiendo comparar flujos continuos de datos contra el perfil del usuario con mínima fricción.

### 1.0.3. Objetivos y Desafíos

El sistema busca resolver desafíos críticos identificados en el estado del arte, como la **fragmentación de modalidades** y la necesidad de **gene-**

**ralización** en escenarios reales. El objetivo final es alcanzar tasas de error competitivas —con un *Equal Error Rate* (EER) en dispositivos móviles— utilizando ráfagas cortas de actividad (apenas 50 caracteres), garantizando así una revocación de acceso casi instantánea ante un posible impostor sin requerir interacción explícita del usuario.

# Capítulo 2

# Estado del Arte

El estado del arte en la biometría del comportamiento en dispositivos móviles abarca una amplia gama de técnicas y metodologías que buscan identificar y autenticar a los usuarios basándose en sus patrones de interacción con el dispositivo. A continuación, se presentan algunas de las investigaciones y desarrollos más relevantes en este campo.

## 2.1. IJCB 2022 MobileB2C: Competición de Autenticación Basada en Comportamiento Móvil

El trabajo de Stragapede et al. 2022*b* presenta una **evaluación comparativa** de sistemas de autenticación móvil basados en **biometría conductual**, capturada de manera transparente mientras el usuario interactúa con su dispositivo.

### Base de Datos Utilizada

El estudio emplea la base pública **BehavePassDB**, recopilada en condiciones reales, que incluye:

- Dinámica de tecleo (*keystroke*)
- Lectura de texto (*text reading*)
- Deslizamiento de galería (*gallery swiping*)
- Toques en pantalla (*tapping*)
- Sensores como acelerómetro, giroscopio, magnetómetro, entre otros

Para el *benchmarking*, además de la identeficación del usuario legítimo, se consideran dos tipos de ataques de impostores:

- **Impostores aleatorios**: otro usuario con un dispositivo distinto.
- **Impostores hábiles**: individuos que intentan imitar al usuario legítimo.

### Conclusiones

Los resultados de la competición **MobileB2C** muestran que la autenticación basada en comportamiento es **viable**, aunque sigue siendo un desafío complejo debido a la variabilidad del entorno y la dificultad de modelar múltiples modalidades de interacción.

Como contribución adicional, el estudio consolida **MobileB2C como una competición continua**, proporcionando una base de datos abierta y un protocolo estándar que facilita nuevas investigaciones en autenticación conductual bajo condiciones realistas.

Cuadro 2.1: Resultados durante la fase de evaluación (AUC [%])

| # | Team | Mixed | Random | Skilled |
|---|---|---|---|---|
| **Task 1: Keystroke** | | | | |
| 1 | NUS-UoA-UoM | **66.37** | **64.77** | **67.91** |
| 2 | HCI Essen | 51.12 | 53.02 | 51.23 |
| 3 | HBKU CS Lab | 51.25 | 49.38 | 53.13 |
| 4 | JAIRG | 45.57 | 52.29 | 39.89 |
| **Task 2: Text Reading** | | | | |
| 1 | HCI Essen | **57.63** | **61.27** | **53.98** |
| 2 | NUS-UoA-UoM | 54.89 | 58.49 | 51.29 |
| 3 | JAIRG | 50.63 | 50.00 | 41.25 |
| 4 | HBKU CS Lab | 48.27 | 59.42 | 37.13 |
| **Task 3: Gallery Swiping** | | | | |
| 1 | HBKU CS Lab | **61.54** | **67.35** | **55.73** |
| 2 | JAIRG | 55.94 | 61.95 | 50.62 |
| 3 | NUS-UoA-UoM | 55.66 | 55.54 | 55.77 |
| 4 | HCI Essen | 54.72 | 57.30 | 51.17 |
| **Task 4: Tapping** | | | | |
| 1 | HBKU CS Lab | **59.58** | **57.22** | **61.94** |
| 2 | NUS-UoA-UoM | 52.39 | 54.72 | 50.06 |
| 3 | JAIRG | 46.25 | 48.75 | 43.75 |
| 4 | HCI Essen | 43.89 | 40.16 | 47.62 |

## 2.2. BehaveFormer: A Framework with Spatio-Temporal Dual Attention Transformers for IMU enhanced Keystroke Dynamics

El trabajo de Senerath et al. 2023 propone *BehaveFormer*, un sistema de autenticación continua en dispositivos móviles basado en la biometría del

comportamiento. El enfoque combina dinámicas de tecleo con datos procedentes de sensores inerciales (IMU), presentes de forma estándar en la mayoría de teléfonos inteligentes. El núcleo del modelo es el *Spatio-Temporal Dual Attention Transformer* (STDAT), una arquitectura basada en *transformers* que emplea mecanismos de atención tanto temporal como de canal para capturar patrones característicos en el comportamiento del usuario.

Para la modalidad de tecleo, se utilizan secuencias de di-gramas y trigramas enriquecidas con tiempos de pulsación (*hold*), transiciones entre eventos y latencias entre teclas. En el caso de la IMU, se extraen derivadas de primer y segundo orden sobre los ejes tridimensionales y se aplica la transformada rápida de Fourier (FFT), obteniendo un vector descriptivo de 36 características por instante.

Cada modalidad es procesada por un STDAT independiente; posteriormente ambas representaciones se fusionan mediante concatenación para generar una incrustación final del usuario. El entrenamiento se realiza mediante *triplet loss*, favoreciendo que muestras del mismo usuario queden próximas en el espacio de representación y que las de distintos usuarios estén separadas.

El método se evalúa sobre tres conjuntos de datos ampliamente utilizados en autenticación continua: Aalto DB, HMOG DB y HuMIdb. Los resultados muestran mejoras significativas frente al estado del arte previo, alcanzando por ejemplo un EER del 1.80 % utilizando solo tecleo en Aalto DB, y un EER del 2.95 % al combinar tecleo e IMU en HuMIdb. Estas cifras confirman la eficacia de la fusión multimodal y del mecanismo de atención dual.

En conjunto, BehaveFormer demuestra que la combinación de información de tecleo y sensores inerciales, unida a arquitecturas basadas en transformers, constituye una vía sólida para sistemas de autenticación pasiva y continua, incrementando la seguridad sin exigir interacción explícita del usuario.

## 2.3. Type2Branch: Keystroke Biometrics based on a Dual-branch Architecture with Attention Mechanisms and Set2set Loss

En esta sección se analiza el artículo de González et al. 2025 que introduce *Type2Branch*, un sistema de autenticación basado en la dinámica de tecleo que se destaca por su capacidad de generalización en escenarios reales, superando las limitaciones de modelos anteriores como *TypeNet*.

### 2.3.1. Evolución de las Arquitecturas de Aprendizaje Profundo

Históricamente, los sistemas de *Keystroke Dynamics* se basaban en redes recurrentes simples para modelar secuencias temporales. Sin embargo, las limitaciones de estas para capturar patrones locales llevaron al desarrollo de enfoques híbridos:

- **Arquitectura Dual-Branch:** El modelo propone una rama de Redes Neuronales Convolucionales (CNN) para la extracción de características locales y una rama de Redes Recurrentes (RNN/LSTM) para las dependencias temporales globales.
- **Mecanismos de Atención:** A diferencia de los modelos precedentes como *TypeNet*, se introducen capas de atención que permiten al sistema ponderar eventos de pulsación específicos que poseen mayor carga discriminativa para la identidad del usuario.

### 2.3.2. Funciones de Pérdida y Verificación

Uno de los mayores avances del artículo es la transición de funciones de pérdida punto a punto hacia enfoques de conjuntos:

- **Set2set Loss:** Frente al tradicional *Triplet Loss*, la función *Set2set* permite comparar un conjunto de muestras de enrolamiento contra una muestra de consulta. Esto es crítico para la **autenticación pasiva**, donde la decisión se basa en el flujo continuo de datos y no en una única entrada estática.

### 2.3.3. Métricas de Rendimiento en Autenticación Pasiva

El rendimiento reportado establece un nuevo estándar para escenarios de texto libre (*free-text*) y dispositivos heterogéneos:

Cuadro 2.2: Rendimiento de Type2Branch en entornos reales.

| **Escenario** | **Usuarios** | **EER (%)** | **Longitud de Secuencia** |
|---|---|---|---|
| Desktop (Escritorio) | 15,000 | 0,77 % | 50 caracteres |
| Mobile (Táctil) | 5,000 | 1,03 % | 50 caracteres |

### 2.3.4. Importancia para la Autenticación Pasiva

El modelo demuestra que es posible alcanzar una alta tasa de precisión con ráfagas cortas de actividad (50 caracteres). En el contexto de la seguridad continua, esto reduce el *Time-to-Detection* (TTD) de un posible impostor, permitiendo una revocación de acceso casi instantánea sin intervención del usuario.

## 2.4. Estado del Arte: Autenticación Continua y Ensamblados de Redes Profundas

El estudio de la biometría de teclado (*Keystroke Dynamics*) ha transitado desde el análisis de textos fijos hacia la autenticación pasiva y continua en entornos de texto libre. El trabajo de Acien and et al. 2022 contextualiza este avance mediante la comparación de técnicas tradicionales frente a

arquitecturas de aprendizaje profundo.

### 2.4.1. Evolución de las Técnicas de Clasificación

La literatura previa se divide fundamentalmente en dos vertientes metodológicas:

- **Aproximaciones Clásicas:** Se basan en algoritmos de aprendizaje automático supervisado como *K-Nearest Neighbors* (KNN), *Naive Bayes* y *Support Vector Machines* (SVM). Aunque eficaces en datasets pequeños, presentan limitaciones de escalabilidad y precisión en escenarios de autenticación continua donde el flujo de datos es masivo y desestructurado.
- **Aprendizaje Profundo (Deep Learning):** El estado del arte reciente destaca el uso de Redes Neuronales Convolucionales (CNN) para extraer características espaciales de las pulsaciones y Redes Neuronales Recurrentes (RNN), específicamente LSTM, para capturar la dependencia temporal. Aversano et al. introducen el concepto de *Ensemble Learning*, combinando múltiples clasificadores base para reducir la varianza y mejorar la robustez del sistema.

### 2.4.2. Extracción de Características y Representación de Datos

El consenso en las investigaciones actuales identifica tres métricas temporales críticas para la biometría conductual:

1. **Dwell Time (DT):** El tiempo que una tecla permanece presionada.
2. **Flight Time (FT):** El intervalo entre la liberación de una tecla y la pulsación de la siguiente.
3. **Inter-key Interval (IKI):** El tiempo transcurrido entre dos pulsaciones consecutivas.

El artículo subraya que la integración de estos rasgos en arquitecturas profundas permite omitir la ingeniería de características manual, ya que la red aprende representaciones jerárquicas de los patrones de tecleo.

### 2.4.3. Desafíos en la Autenticación Pasiva

A pesar de los avances, el estado del arte identifica brechas significativas que este estudio busca solventar:

- **Fragmentación de Datos:** La mayoría de los estudios previos utilizan datasets pequeños (menos de 100 usuarios). El artículo destaca la necesidad de *benchmarks* masivos, proponiendo un dataset integrado de más de 160,000 usuarios.
- **Generalización:** La dificultad de los modelos para mantener la precisión cuando el usuario cambia de contexto de escritura o de dispositivo.
- **Toma de Decisión Continua:** La transición de una autenticación de un solo paso (login) a una monitorización constante sin fricción para el usuario.

## Estado del Arte: Biometría Conductual en Dispositivos Móviles

El artículo de Stragapede et al. 2022*a* sitúa su investigación en el contexto de la autenticación pasiva mediante biometría conductual en smartphones, analizando tanto la interacción táctil como los sensores de movimiento (background sensors).

### Clasificación de Técnicas Anteriores

Los autores clasifican los trabajos previos principalmente en base a las modalidades biométricas y las arquitecturas de procesamiento:

- **Dinámica de Tecleo (Keystroke Dynamics):** Se divide en escenarios de *texto fijo* (contraseñas) y *texto libre*. Se mencionan aproximaciones basadas en redes neuronales recurrentes, específicamente Long Short-Term Memory (LSTM) para autenticación a gran escala.
- **Gestos Táctiles (Touch Gestures):** Técnicas que analizan el desplazamiento (scrolling), toques (tapping) y dibujo de patrones.
- **Sensores de Movimiento (Background Sensors):** Uso de acelerómetros, giroscopios y magnetómetros para capturar patrones físicos del usuario durante la interacción.
- **Arquitecturas de Aprendizaje:** El artículo distingue entre modelos estadísticos tradicionales y enfoques modernos de Deep Learning (como CNNs para gestos y LSTMs para secuencias temporales).

## Limitaciones Identificadas

El estudio subraya varias deficiencias en la literatura actual:

1. **Fragmentación de Modalidades:** La mayoría de los estudios se centran en una sola modalidad (solo tecleo o solo acelerómetro), perdiendo la sinergia de la fusión multimodal.
2. **Escalabilidad y Realismo:** Muchos trabajos utilizan bases de datos pequeñas o recolectadas en entornos de laboratorio controlados que no reflejan el uso cotidiano (ïn-the-wild").
3. **Ventanas de Tiempo Elevadas:** Algunos sistemas requieren periodos de observación demasiado largos para alcanzar precisiones aceptables, lo que reduce la eficacia de la autenticación pasiva inmediata.

## Problema Específico a Resolver

El vacío que este artículo intenta llenar es la **evaluación exhaustiva de la fusión de múltiples biometrías conductuales** (tecleo, scroll, dibujo, sensores de fondo) bajo un protocolo común y utilizando una base de datos

pública de gran escala (*HuMIdb*). El objetivo es determinar cuál es la combinación mínima de sensores y tiempo de análisis necesaria para lograr una autenticación robusta y transparente.

# Bibliografía

Acien, A. y et al.: 2022, *Expert Systems with Applications*

González, N., Stragapede, G., Vera-Rodriguez, R., y Tolosana, R.: 2025, *Type2Branch: Keystroke Biometrics based on a Dual-branch Architecture with Attention Mechanisms and Set2set Loss*

Senerath, D., Tharinda, S., Vishwajith, M., Rasnayaka, S., Wickramanayake, S., y Meedeniya, D.: 2023, *BehaveFormer: A Framework with Spatio-Temporal Dual Attention Transformers for IMU enhanced Keystroke Dynamics*

Stragapede, G., Vera-Rodriguez, R., Tolosana, R., Morales, A., Acien, A., y Le Lan, G.: 2022a, *Pattern Recognition Letters* **157**, 35

Stragapede, G., Vera-Rodriguez, R., Tolosana, R., Morales, A., Fierrez, J., Ortega-Garcia, J., Rasnayaka, S., Seneviratne, S., Dissanayake, V., Liebers, J., Islam, A., Belhaouari, S. B., Ahmad, S., y Jabin, S.: 2022b, *IJCB 2022 Mobile Behavioral Biometrics Competition (MobileB2C)*