UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

TESIS DOCTORAL

Autenticación Continua en Dispositivos Móviles Basada en Inteligencia Artificial

D. Juan Manuel Espín López

2024

UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

TESIS DOCTORAL

Autenticación Continua en Dispositivos Móviles Basada en Inteligencia Artificial

Autor: D. Juan Manuel Espín López

Director/es: D. Javier Marín-Blázquez Gómez

D. Francisco Esquembre Martínez

D. Alberto Huertas Celdrán

UNIVERSIDAD DE MURCIA

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA EN MODALIDAD DE COMPENDIO O ARTÍCULOS PARA OBTENER EL TÍTULO DE DOCTOR

Aprobado por la Comisión General de Doctorado el 19-10-2022

D./Dña. Juan Manuel Espín López

doctorando del Programa de Doctorado en

Informática

de la Escuela Internacional de Doctorado de la Universidad Murcia, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

Autenticación Continua en Dispositivos Móviles Basada en Inteligencia Artificial

y dirigida por,

D./Dña. Javier G. Marín-Blázquez

D./Dña. Francisco Esquembre Martínez

D./Dña. Alberto Huertas Celdrán

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita, cuando se han utilizado sus resultados o publicaciones.

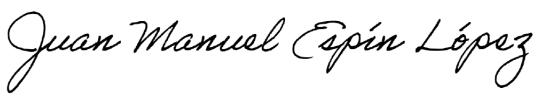
Además, al haber sido autorizada como compendio de publicaciones o, tal y como prevé el artículo 29.8 del reglamento, cuenta con:

- *La aceptación por escrito de los coautores de las publicaciones de que el doctorando las presente como parte de la tesis.*
- *En su caso, la renuncia por escrito de los coautores no doctores de dichos trabajos a presentarlos como parte de otras tesis doctorales en la Universidad de Murcia o en cualquier otra universidad.*

Del mismo modo, asumo ante la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada, en caso de plagio, de conformidad con el ordenamiento jurídico vigente.

En Murcia, a 06 de febrero de 2024

Fdo.: Juan Manuel Espín López



Esta DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD debe ser insertada en la primera página de la tesis presentada para la obtención del título de Doctor.

Información básica sobre protección de sus datos personales aportados	
Responsable:	Universidad de Murcia. Avenida teniente Flomesta, 5. Edificio de la Convalecencia. 30003; Murcia. Delegado de Protección de Datos: dpd@um.es
Legitimación:	La Universidad de Murcia se encuentra legitimada para el tratamiento de sus datos por ser necesario para el cumplimiento de una obligación legal aplicable al responsable del tratamiento. art. 6.1.c) del Reglamento General de Protección de Datos
Finalidad:	Gestionar su declaración de autoría y originalidad
Destinatarios:	No se prevén comunicaciones de datos
Derechos:	Los interesados pueden ejercer sus derechos de acceso, rectificación, cancelación, oposición, limitación del tratamiento, olvido y portabilidad a través del procedimiento establecido a tal efecto en el Registro Electrónico o mediante la presentación de la correspondiente solicitud en las Oficinas de Asistencia en Materia de Registro de la Universidad de Murcia

La siguiente Tesis Doctoral es una recopilación de los siguientes artículos publicados, siendo el doctorando el autor principal en todos ellos:

- Juan Manuel Espín López, Alberto Huertas Celdrán, Javier G. Marín-Blázquez, Francisco Esquembre, Gregorio Martínez Pérez. "**S3: An AI-Enabled User Continuous Authentication for Smartphones Based on Sensors, Statistics and Speaker Information**", *Sensors*, vol 21 (11), pp 3765, 2021.
  DOI: https://doi.org/10.3390/s21113765
  JIF 2020: 3.576 (Q1)
  JIF 2021: 3.847 (Q2)

- Juan Manuel Espín López, Alberto Huertas Celdrán, Francisco Esquembre, Gregorio Martínez Pérez, Javier. G. Marín-Blázquez, "**A Supervised ML Biometric Continuous Authentication System for Industry 4.0**", *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 9132-9140, Dec. 2022.
  DOI: 10.1109/TII.2022.3171321.
  JIF 2021: 11.648 (D1)
  JIF 2022: 12.300 (D1)

- Juan Manuel Espín López, Alberto Huertas Celdrán, Francisco Esquembre, Gregorio Martínez Pérez, Javier. G. Marín-Blázquez, "**CGAPP: A Continuous Group Authentication Privacy-Preserving Platform for Industrial Scene**", *Journal of Information Security and Applications*, vol. 78, pp. 103622, Nov. 2023.
  DOI: 10.1016/j.jisa.2023.103622
  JIF 2021: 4.960 (Q2)
  JIF 2022: 5.600 (Q2)

Contenido

Agradecimientos **iii**

Resumen **v**

I Introducción y motivación v
II Objetivos x
III Metodología x
IV Resultados xii
V Conclusiones y trabajo futuro xv

Bibliografía **xx**

Publicaciones que componen la Tesis Doctoral (PhD Tesis)

1 S3: An AI-Enabled User Continuous Authentication for Smartphones Based on Sensors, Statistics and Speaker Information **3**

2 A Supervised ML Biometric Continuous Authentication System for Industry 4.0 **5**

3 CGAPP: A Continuous Group Authentication Privacy-Preserving Platform for Industrial Scene **7**

Agradecimientos

Quiero dedicar dedicar este apartado para expresar mi agradecimiento a todas las personas que han contribuido (o se han visto afectadas) de manera significativa en la realización de esta tesis doctoral. A lo largo de este viaje académico, me he enfrentado a desafíos y obstáculos, pero también he experimentado momentos de crecimiento personal y logros profesionales que han dado forma a mi trayectoria académica y mi vida personal.

A mi Anita, quiero agradecerle de todo corazón su apoyo incondicional, su paciencia infinita y su comprensión durante este largo periodo. Tuvimos que sacrificar momentos de nuestra vida juntos para dedicarme al estudio y la investigación, pero siempre estuviste a mi lado, brindándome tu aliento y motivación. Sin tu amor y respaldo, este logro no hubiera sido posible.

A mis padres les agradezco que desde pequeño me inculcaran valores de perseverancia, resiliencia y superación personal. Su amor incondicional y su dedicación hacia mi educación han sido fundamentales en mi camino hacia este título de doctorado. La confianza de mi familia en mis capacidades y su constante estímulo han sido motores que me impulsaron a alcanzar mis metas. A mis abuelos, que mil veces me preguntaron cuando tenía el examen del doctorado y mil veces me decían que largo se me estaba haciendo.

A mis amigos, aquellos con los que compartía momentos de alegría y llegada una hora me retiraba, o había llegado más tarde, por el trabajo dedicado a esta tesis. También a Alejandro y Jesús por ser una parte vital de mi vida académica. También a todos aquellos que os habéis cruzado, os he contado mis ideas, dudas, y hemos tenido discusiones enriquecedoras que fomentaron mi aprendizaje, allanaron el camino y me motivaron a esforzarme aún más.

A mis directores de tesis, Paco, Javi y Alberto, les dedico unas palabras de agradecimiento sincero por su contribución a este trabajo.

Esquembre, quiero agradecerte por contar conmigo para aplicar lo aprendido y aprender un nuevo mundo. Además, por abrir mi perspectiva más allá de las teorías matemáticas y mostrarme cómo aplicarlas en el mundo real, enjambres, difusión de enfermedades, entre otros. Tu enfoque en la aplicación práctica de los conceptos me ha permitido comprender la importancia y la utilidad de la investigación científica.

Javi, te estoy enormemente agradecido por introducirme en el apasionante mundo de la inteligencia artificial y por permitirme tener una caja de herramientas variada y enseñarme que no todos los tornillos son iguales. Tu conocimiento experto y tus habilidades técnicas me han capacitado para abordar problemas complejos y aplicar soluciones innovadoras. Aunque no olvidaré los micro-infartos provocados por los "¡¡Paren las máquinas!!"

Alberto, mi gratitud hacia ti es infinita. Sin tu aparición en el proyecto, perseverancia, dedicación, trabajo incansable, ideas brillantes y propuestas valiosas, esta tesis no habría sido posible. Tu compromiso con mi crecimiento académico y tu capacidad para desafiar mis límites me han empujado a alcanzar metas que nunca imaginé. Tu mentoría y dirección han sido esenciales en cada etapa de esta investigación, y estoy eternamente agradecido por tu apoyo incondicional.

Por último, quiero agradecer a Gregorio Martínez por guiarme en un momento en el que consideraba rendirme. Tu idea innovadora y la línea de investigación que propusiste fueron un verdadero punto de inflexión en este proyecto. Agradezco especialmente el apoyo brindado por todo tu grupo de investigación, participativos y abiertos desde el momento cero. La ayuda que me han otorgado es incalculable.

A todos vosotros, gracias! Cada trocito de esta tesis os pertenece.

Resumen

I Introducción y motivación

Los teléfonos inteligentes (*smartphones*) se han convertido con el paso de los años en herramientas indispensables para las personas y las empresas. Muy lejos quedaron aquellos años en los que el teléfono móvil solo permitía hablar y enviar o recibir unos limitados SMS. Actualmente, el teléfono móvil se utiliza para multitud de actividades, desde realizar tareas del trabajo (leer correos, preparar presentaciones o atender videoconferencias, entre otras) hasta pasar tiempo libre (viendo alguna serie, realizando *scroll* en alguna red social o comprando y vendiendo artículos de segunda mano).

Para poder tener en la palma de nuestra mano todas las funcionalidades que nos ofrecen, estos dispositivos deben almacenar una ingente cantidad de información sobre nosotros mismos. Esta información incluye habitualmente datos sensibles y que afectan a la privacidad del usuario, como por ejemplo, la ubicación a lo largo del día, información bancaria, familiar, sobre la salud, etc. Una vez obtenido desbloqueado un dispositivo, el acceso a la información es sencillo. Por ello, es crítico proteger el acceso a estos dispositivos para que solo puedan acceder los usuarios autorizados. La autenticación es el proceso mediante el cuál se verifica la identidad de los usuarios para darles el acceso correspondiente a su nivel de autorización.

Blindar la obtención de los datos almacenados en los dispositivos móviles, no solo smartphones, sino también tabletas y ordenadores portátiles (*laptops*) ha sido objeto de investigación y trabajo constante de compañías desarrolladoras de software, concretamente de sistemas operativos (SO). Entre los sistemas de autenticación y acceso, destacan los más conocidos por los usuarios: patrones, contraseñas, códigos e incluso los biométricos, como la huella dactilar o el reconocimiento facial (muy presente en la generación actual de dispositivos móviles) [1]. Todos los sistemas de acceso implantados mediante biometría, presentan un sistema de acceso paralelo basado en códigos, debido a que, en diversas ocasiones, están tecnologías requieren de ciertas condiciones ambientales para poder operar. En todos estos casos, el esfuerzo se centra en evitar el acceso no autorizado al dispositivo, pero una vez son vulnerados, no hay implantadas herramientas para bloquear el ataque y evitar problemas mayores. Por ejemplo, una persona que consiga acceder a un dispositivo móvil podría tener acceso a una variedad de aplicaciones que siempre permanecen abiertas en el dispositivo.

Pero no solo nos debe preocupar la información a la que pueden acceder los atacantes, sino también las tareas, órdenes o acciones que pueden realizar en nuestro nombre. Por

ejemplo, atacantes o personas sin autorización podrían efectuar una llamada maliciosa o enviar un email fraudulento a nuestro banco, administrador de fincas, negocios, dinero, o incluso a nuestra familia y realizar un perjuicio mayor. De hecho, existe incluso una acción potencialmente delicada para la que una persona ajena no necesita siquiera vulnerar el acceso: *responder* una llamada. Una llamada puede ser respondida por cualquier persona, sin necesidad de ser el propietario del teléfono al que va dirigida la llamada. (Para mostrar el potencial peligro, consideremos el caso de un ladrón que substrae el dispositivo de un comercial, el cual recibe una llamada para confirmar una transacción, y el ladrón solicita que se haga un cambio de cuenta bancaria y consigue substraer el dinero de la operación.) En definitiva, el mayor abanico de usos personales o profesionales de los dispositivos móviles hace que sea también cada vez mayor la variedad de situaciones en las que se hace imprescindible algún mecanismo de autenticación en nuestros dispositivos.

Autenticación en la Industria 4.0

Como se menciona anteriormente, estos dispositivos (smartphones, tabletas o laptops) no son solo de uso personal. En particular, como parte de sus usos profesionales, muchas de las líneas de producción de la actual Industria 4.0 están automatizadas y controladas mediante dispositivos electrónicos que, a su vez, son operados o programados por empleados altamente capacitados. Dado que los procesos están orientados a maximizar la producción, cualquier operación incorrecta o dañina puede causar pérdidas significativas de tiempo y/o dinero y afectar seriamente la producción. Esto hace que la seguridad sea una cuestión central en las fábricas modernas [2]. La seguridad comienza con la autenticación de los trabajadores [3] que operan las máquinas de producción. Permitir que un trabajador realice una tarea que no está dentro de sus deberes prescritos o para la cual aún no ha recibido capacitación puede representar un riesgo significativo de fracaso. Peor aún, permitir el acceso a la fábrica o a sus sistemas a personas no autorizadas o atacantes (saboteadores) es un grave riesgo de seguridad siempre presente.

La autenticación es un tema amplio y bien estudiado que aún plantea preguntas de investigación abiertas. Aunque algunas investigaciones en autenticación están considerando el uso de algunas técnicas futuristas, como el muestreo de ADN, pero los mecanismos más comúnmente utilizados en las industrias para la autenticación de usuarios son las tarjetas de identificación, las contraseñas y la biometría [4]. Las tarjetas de identificación permiten un acceso rápido a las instalaciones a través de puertas o torniquetes, pero pueden perderse, robarse, clonarse o intercambiarse entre los trabajadores. Las contraseñas pueden proporcionar un acceso moderadamente rápido a la fábrica y a los sistemas informáticos, pero también pueden ser robadas o intercambiadas (y frecuentemente olvidadas). Finalmente, los sistemas biométricos, como el reconocimiento facial o el de huellas dactilares, son mecanismos de acceso de alta velocidad que son más difíciles de suplantar y, definitivamente, no olvidables. Aún así, su uso requiere sensores específicos, que a veces pueden interferir con el equipo de protección personal que usan los trabajadores, como gafas o mascarillas, haciendo que su uso sea poco práctico o incluso imposible en algunas situaciones. Por lo tanto, para estas situaciones es necesario un método de autenticación más pasivo, no intrusivo y que no restrinja el uso de equipamiento.

Autenticación Continua

Una manera efectiva de combatir los problemas de seguridad mencionados anteriormente en los dispositivos móviles es agregar a estos dispositivos un sistema de *Autenticación Continua* [5]. Estos sistemas tienen como objetivo principal confirmar la identidad del

usuario de forma constante, lo que, a su vez, mejora la usabilidad de los dispositivos [6]. En concreto, estos sistemas pueden analizar, entre otros factores [7], el comportamiento de los usuarios en el dispositivo. Para ello, elaboran un modelo de reconocimiento del usuario basado en datos procedentes de diversas fuentes, entre las que se incluyen los sensores del teléfono, las pulsaciones de teclas, las estadísticas de uso de aplicaciones e incluso los patrones de movimiento de los dedos sobre la pantalla del dispositivo. Una vez que se ha creado un perfil de comportamiento para un usuario específico, se van extrayendo nuevos datos de forma continua en el tiempo, que son evaluados por el modelo para obtener una puntuación de similitud. En el caso de que dicha puntuación no alcance los valores necesarios, se considera que la muestra no pertenece al usuario legítimo y, por tanto, el dispositivo está siendo utilizado por otra persona.

Un dispositivo móvil ofrece diversidad de fuentes con las que poder trabajar y construir un sistema de autenticación continua. Desde los datos de sensores intrínsecos al dispositivo (el acelerómetro, el giróscopo o el magnetómetro [8, 9]), pasando por datos dinámicos (pulsación de teclas, deslizamiento de dedos por la pantalla [10, 11]), datos del entorno (red WiFi, localización GPS, o dispositivos Bluetooth conectados), estadísticos de uso de aplicaciones o también información biométrica (facial, huella dactilar [12], iris [13, 14], voz [15], o incluso el paso [16]). Aunque algunas soluciones pueden utilizar una única fuente, la mayoría de los trabajos suelen usar una combinación de varias de ellas [17, 18]. La combinación de fuentes puede dotar de una mayor robustez al sistema y ofrecer una mayor disponibilidad de datos.

Los sistemas de autenticación continua usan más frecuentemente técnicas de Inteligencia Artificial (IA), de Aprendizaje Automático o *Machine Learning* (ML) o de Aprendizaje Profundo o Deep Learning (DL), que otros enfoques diferentes, como los sistemas basados en reglas, estadísticos, conocimiento, etc. [19, 20]. Estas técnicas de IA se componen de dos partes bien diferenciadas: el entrenamiento del modelo y el test o evaluación. Durante el entrenamiento de los modelos, se utilizan datos para adaptar los pesos de los modelos de inteligencia artificial y que estos aprendan las distribuciones o tareas que se requiere que aprendan. Para ello, se suele utilizar una función de aprendizaje u optimizador, siendo el más utilizado el SGD (Stocastic Gradient Descent) o alguna de sus variantes, que va adaptando los pesos del modelo. Tras varias iteraciones, se evalúa la función de pérdida o de precisión y se detiene, en caso satisfactorio, el entrenamiento. Una vez que el modelo ha sido entrenado, cuando aparecen nuevos datos, estos pueden ser evaluados por el modelo, produciendo así una respuesta en función de la distribución que ha aprendido.

Las diferentes técnicas de IA se pueden agrupar en dos enfoques diferentes atendiendo a si utilizan datos etiquetados o no: enfoque supervisado y enfoque no supervisado. Los métodos no supervisados son útiles cuando no se tiene acceso a un conjunto de datos etiquetado o cuando se necesita detectar patrones inusuales en el comportamiento del usuario (enfoque de detección de valores atípicos o *outliers*). En este enfoque, los sistemas solo son entrenados con los datos o información del usuario que pretenden identificar, generando un sistema que modele su comportamiento y detecte cuando no es el usuario indicado. Un ejemplo sería el uso de algoritmos de agrupamiento o *clustering*, como el denominado *K-means* [21], para agrupar diferentes comportamientos del usuario y luego detectar anomalías que no pertenezcan a esos grupos [20]. Esto es útil en situaciones donde se necesita una detección temprana de amenazas sin la necesidad de datos etiquetados, como en la detección de intrusiones en sistemas de seguridad [22].

Por otro lado, los métodos supervisados son apropiados cuando se dispone de un conjunto de datos etiquetado con ejemplos de autenticación exitosa y fallida. Estos algoritmos utilizan estos ejemplos para aprender patrones y características que ayuden a distinguir

entre usuarios legítimos y posibles impostores. Ejemplos de métodos supervisados incluyen el aprendizaje automático con clasificadores como Support Vector Machines (SVM) [23], Random Forests [24] o Redes Neuronales [25], donde se utilizan datos históricos para entrenar el modelo y luego se evalúa su desempeño en tiempo real. Esto es útil en situaciones donde se requiere un alto nivel de precisión y se pueden disponer de datos etiquetados [18]. Resulta interesante analizar y evaluar las diferencias de precisión y robustez de estos enfoques en tareas y contextos donde nunca antes habían sido estudiados.

Autenticación de Grupo

Independientemente del enfoque seguido, la mayoría de los sistemas de autenticación tradicionales se centran en la identificación individual, donde se verifica si un dispositivo es utilizado por un usuario legítimo específico. Por lo general, para crear modelos efectivos de autenticación individual, se requiere acceso a los datos personales y privados de esos usuarios. A medida que las preocupaciones sobre la privacidad de los individuos aumentan, incluso en entornos laborales, se ha generado un interés creciente en sistemas de autenticación que priorizan la protección de la privacidad.

En algunos escenarios, sin embargo, no es necesario identificar individualmente a los usuarios; es suficiente con comprobar que pertenecen a un grupo determinado (con la debida autorización). En este contexto, la denominada *autenticación de grupo* [26] se presenta como una solución adecuada. Estos sistemas verifican si un usuario es miembro de un grupo específico respetando su privacidad. Para garantizar la privacidad de los usuarios, estos sistemas deben permitir que los usuarios indiquen su pertenencia al grupo sin revelar datos personales privados. Los miembros del grupo deben tener los mismos privilegios, como acceso a ubicaciones, responsabilidades y tareas. La autenticación de grupo mejora la privacidad al confirmar solo la pertenencia al grupo sin exponer datos personales privados. La combinación de un sistema de autenticación continua junto con la autenticación grupal nunca ha sido propuesto, ni estudiado o evaluado en algún trabajo anterior al que aquí se presenta.

Aprendizaje Federado

Hasta hace poco, muchos algoritmos de aprendizaje para desarrollar modelos de autenticación requerían que los datos personales de los usuarios se extrajeran de los dispositivos y se enviaran a un servidor o repositorio, lo que planteaba preocupaciones de privacidad. Sin embargo, en el año 2016 surge un nuevo enfoque llamado Aprendizaje Federado (FL), propuesto por Google [27], que elimina la necesidad de compartir datos privados. El enfoque FL permite construir modelos utilizando la contribución de múltiples participantes sin que estos compartan directamente sus datos. En cambio, cada participante federado comparte información como pesos o gradientes de entrenamiento. Luego, esta información se combina en un modelo global único, que se distribuye nuevamente a cada participante. Después de varias iteraciones de este proceso, se obtiene un modelo final entrenado con el conocimiento de todos los participantes. Este enfoque de entrenamiento aborda de manera efectiva el problema de la exposición de datos y la privacidad que preocupa a muchos usuarios. En este caso, los datos privados de los usuarios ya no necesitan salir de sus dispositivos para construir el modelo, lo que mejora significativamente la protección de la privacidad.

A pesar de la mejora en la protección de la privacidad, el enfoque de entrenamiento mencionado aún puede ser vulnerable a diversos tipos de ataques, como se señala en [28]. Los ataques pueden ser agrupados en dos grandes conjuntos, *Model Poisoning* [29], cuando

el ataque se centra en el modelo, o *Data Poisoning* [30], si los ataques se centran en los datos. Pero, en ambos conjuntos de ataques, el objetivo es el mismo: afectar al rendimiento del modelo. Además, se puede atender al lugar en el que se cometen estos ataques. Por ejemplo, en una arquitectura cliente-servidor, el ataque puede ser llevado a cabo desde ambos emplazamientos. Sin embargo, en muchos escenarios, dado que el servidor es proporcionado por la parte que tiene interés en protegerlo y que no pueda ser atacado, éste se considera honesto y confiable, lo que significa que los ataques adversariales generalmente solo pueden ocurrir en el lado del cliente.

Los ataques a un sistema federado pueden variar en complejidad, desde simples hasta muy elaborados. Pero los ataques más preocupantes son aquellos que pueden llevarse a cabo sin requerir un conocimiento técnico profundo de la aplicación. Dos tipos de ataques destacan en este contexto: la inyección [31] y la perturbación de datos [32], que son ataques de tipo Data Poisoning y que pueden ser llevados a cabo en la parte del cliente. En un ataque de inyección, un usuario malicioso puede utilizar el dispositivo de un trabajador legítimo para introducir datos de comportamiento falsos o suplantados en el sistema. Por otro lado, en un ataque de perturbación de datos, un trabajador comprometido altera intencionalmente su propio comportamiento para provocar fallos en el sistema.

Estos tipos de ataques representan un desafío importante para la seguridad de los sistemas de aprendizaje federado, ya que pueden socavar la integridad de los datos y poner en riesgo la privacidad de los usuarios. Por lo tanto, la protección contra estos tipos de amenazas sigue siendo una preocupación importante en la implementación de sistemas de autenticación y aprendizaje federado.

Preguntas de Investigación

Esta tesis doctoral se centra en investigar un sistema de autenticación continua que mejore la seguridad de los usuarios de los dispositivos móviles y mejore la privacidad de los datos y la identidad del usuario. Teniendo en cuenta toda la información dada y la dirección de la tesis, surgen varias preguntas de investigación, que guiarán el proceso de investigación de esta tesis doctoral. Las preguntas de investigación son las siguientes:

- PI1: En un escenario de autenticación continua basado en información de sensores, estadísticos y voz, ¿cuál es la dimensión más informativa y cual es la mejor forma de agruparlas en caso de poder usar varias al mismo tiempo?
- PI2: En un escenario de autenticación continua en dispositivos móviles en el que se cuenta con datos etiquetados, ¿cuánto mejoran los sistemas que siguen un enfoque supervisado frente a los no supervisados?
- PI3: En un escenario de autenticación continua en dispositivos móviles, ¿cómo de robustos son los sistemas entrenados de modo supervisado al aparecer nuevos usuarios en el sistema?
- PI4: Para un sistema de autenticación continua grupal basado en el uso de sensores, ¿cuál es el coste en términos de precisión por incrementar la privacidad de los datos?
- PI5: En un sistema de autenticación continua grupal entrenado mediante un esquema federado, ¿cómo de robustos es el sistema frente a los ataques adversariales de tipo inyección de datos y perturbación de datos?
- PI6: En un sistema de autenticación continua grupal entrenado mediante un esquema federado, ¿existen contramedidas que puedan paliar los efectos de los ataques adversariales de tipo inyección y perturbación de datos?

II Objetivos

El objetivo principal de esta tesis doctoral es investigar los sistemas de autenticación continua que mejoren la seguridad de los usuarios de dispositivos móviles y mantengan la privacidad de los datos y la identidad del usuario. Para lograr este objetivo, se plantean las siguientes metas específicas, cada una de las cuales está relacionada con una pregunta de investigación (PI) específica:

- Analizar el estado del arte actual para identificar y definir los pasos necesarios para construir un sistema de autenticación continua en dispositivos móviles. (PI1, PI2, PI4, PI5)
- Creación de un dataset que contenga datos de sensores, estadísticos de uso y voz, de diferentes usuarios utilizando dispositivos móviles. (PI1)
- Evaluar las fuentes de información disponibles para la autenticación continua, como sensores, datos estadísticos de uso y voz. Determinar si es más efectivo combinar estas fuentes de información a nivel de dato o de sistema. (PI1)
- Comparar y evaluar el rendimiento de sistemas de autenticación continua que siguen un enfoque supervisado con aquellos que siguen un enfoque no supervisado, en caso de contar con datos etiquetados. Medir la eficacia y precisión de ambos enfoques en la identificación de usuarios legítimos y la detección de impostores. (PI2)
- Evaluar la robustez de un sistema de autenticación continua entrenado de forma supervisada ante la aparición de nuevos usuarios; es decir, nunca vistos en el entrenamiento.
- Cuantificar el impacto en términos de precisión al aumentar la privacidad de los datos de autenticación, mediante el cambio de paradigma del sistema o mediante la aplicación de entrenamiento federado (FL). Evaluar si los sistemas de autenticación continua pueden proteger la privacidad de los usuarios sin sacrificar la precisión. (PI4)
- Analizar la resiliencia de los sistemas de autenticación continua en un esquema de entrenamiento federado frente a ataques adversariales. Evaluar la degradación del sistema y qué medidas de seguridad son efectivas para mitigar estos ataques. (PI5, PI6)

III Metodología

La presente tesis doctoral se estructura como un compendio de tres artículos científicos publicados en revistas indexadas en el Journal Citation Reports (JCR). Cada uno de estos artículos contribuye de manera significativa a la investigación y responde a preguntas de investigación específicas planteadas anteriormente (Sección I). El enfoque seguido en esta tesis es de naturaleza científica, con una revisión continua del estado del arte, respaldado por una metodología rigurosa y basado en la investigación empírica.

La primera pregunta de investigación (PI1) cuestiona los pasos iniciales para construir un sistema de autenticación continua en dispositivos móviles y la elección de las fuentes de información. Para ello, se evaluaron y estudiaron trabajos previos de autenticación continua en dispositivos móviles, que permitieron elegir las fuentes idóneas para conseguir el objetivo.

En el primer capítulo de este documento, S3: An AI-Enabled User Continuous Authentication for Smartphones Based on Sensors, Statistics and Speaker Information (Article 1–`Sensors`), se presenta la plataforma S3, un sistema de autenticación continua potenciado por inteligencia artificial que combina datos de sensores, estadísticas de aplicaciones y, de manera significativa, la voz del usuario como fuente de información. Mediante experimentos detallados, se analiza la relevancia de cada tipo de dato, se exploran estrategias para su combinación y se determina cuántos días de entrenamiento son necesarios para crear perfiles precisos de los usuarios. Los resultados revelan que la voz desempeña un papel fundamental en la construcción de un sistema de autenticación precisa, superando la importancia de los sensores y las estadísticas de aplicaciones. Además, la combinación de modelos individuales se identifica como la estrategia óptima. Este artículo no solo considera los pasos iniciales y la elección de las fuentes de información, sino que también proporciona datos empíricos sólidos respaldados por un conjunto de datos real. Es importante destacar que una contribución significativa de este trabajo es que se ha puesto dicho conjunto de datos a disposición de la comunidad científica, lo que representa un aporte valioso al campo de la investigación en autenticación continua en dispositivos móviles.

Tras completar los pasos iniciales, obtener y explorar las fuentes de información y conocer la viabilidad de la autenticación continua en dispositivos móviles usando las fuentes propuestas, y tras analizar los resultados obtenidos, se propuso realizar una arquitectura orientada a IA supervisada. No obstante, era necesario, antes de confirmar dicha decisión, obtener la respuesta a la segunda pregunta de investigación (PI2), que constituye uno de los objetivos del capítulo 2 (A Supervised ML Biometric Continuous Authentication System for Industry 4.0 (Article 2–`IEEE_TII`)). Este artículo aborda directamente la PI2 al comparar el rendimiento de un sistema de autenticación continua basado en aprendizaje automático (ML) que sigue un enfoque supervisado frente a uno no supervisado. Dicho sistema fue testeado en un escenario industrial simulado donde, a través de experimentos exhaustivos, se evalúa la precisión y el rendimiento de estos modelos en función de los datos recopilados de sensores, estadísticas de aplicaciones y datos de voz del operador. Los resultados revelan diferencias significativas en el rendimiento, con mejoras notables en los modelos supervisados en comparación con los no supervisados.

El segundo artículo también responde a la PI3, al explorar la robustez de un sistema de autenticación continua entrenado de forma supervisada cuando aparecen nuevos usuarios. En dicho trabajo, se evalúa la generalización de los modelos ante nuevos usuarios desconocidos. Este artículo, al proporcionar datos empíricos sólidos y conclusiones claras en relación con estas preguntas de investigación, contribuye de manera significativa a la comprensión de la autenticación continua entrenada en modo supervisado.

El curso de la tesis doctoral ha evolucionado conjuntamente con las preocupaciones de la sociedad con el objetivo de solventar las dudas y hacer fácil y accesible el acceso a la tecnología. Con el paso del tiempo, los usuarios se han visto muy interesados y preocupados con el acceso a sus datos y su privacidad. Por ello, el tercer capítulo de esta tesis (CGAPP: A Continuous Group Authentication Privacy-Preserving Platform for Industrial Scene (Article 3–`IEEE_JISA`)) se centra en responder la PI4 al investigar el coste en términos de precisión al priorizar la privacidad de los datos en un sistema de autenticación grupal y continuo. Para abordar esta pregunta, se propone una plataforma de autenticación grupal continua en un escenario industrial que emplea aprendizaje federado y garantiza la privacidad de los datos personales de los usuarios. El artículo realiza experimentos para evaluar el rendimiento y la precisión de esta plataforma en comparación con enfoques tradicionales que no se preocupan por la privacidad.

Al usar enfoques federados, hay que tener en cuenta y estudiar el posible efecto que

tienen los ataques adversariales en el sistema de autenticación continua. Por ello, también en este tercer trabajo se analiza cuántos datos maliciosos son necesarios para engañar al modelo y cuánto ruido se requiere para perturbar el sistema de autenticación. Los experimentos revelan la robustez del sistema ante ataques de inyección de datos, mientras que se identifican desafíos significativos en la protección contra ataques de perturbación de datos.

En resumen, para la consecución de esta tesis se ha revisado el estado del arte en autenticación continua, atendiendo a dispositivos móviles, fuentes, enfoques de aplicación, autenticación grupal y entrenamiento federado. Se han llevado a cabo experimentos exhaustivos que han permitido evaluar la precisión, rendimiento y robustez de los sistemas de autenticación desarrollados. Además, se han construido sistemas desplegables en los dispositivos, incluidos los utilizados en entornos industriales. Esta metodología ha permitido investigar y proporcionar respuestas sólidas a preguntas cruciales relacionadas con la autenticación continua en dispositivos móviles, también en el contexto de la Industria 4.0.

IV Resultados

Esta tesis por compendio se compone de tres artículos publicados en revistas científicas indexadas en el índice JCR. A continuación se resumen los resultados obtenidos en cada artículo.

En el primer artículo de esta tesis doctoral, Article 1–`Sensors`, se presenta la Plataforma S3, que recoge datos de sensores, estadísticos de uso de aplicaciones y la voz, que permiten autenticar a los usuarios en las actividades más habituales de uso de un dispositivo móvil. Desde la interacción con aplicaciones hasta el momento de atender una llamada en manos libres. La plataforma se compone de dos partes bien diferenciadas, una aplicación móvil y un servidor. La aplicación móvil solo se encargar de adquirir los datos y comunicarlos al servidor. Es en este, el servidor, donde se lleva a cabo todo el procesamiento: la preparación de los datos, el almacenamiento de los mismos, el entrenamiento de los modelos de comportamiento de los usuarios y el cotejo de nuevas muestras contra estos modelos. Una vez se tiene cotejada una nueva muestra, se envía una respuesta a la aplicación.

Tras analizar el estado del arte y buscar bases de datos públicas que permitieran llevar a cabo la investigación, se comprobó su ausencia y, por tanto, se llevó a cabo una recopilación de datos que dieron lugar a una base de datos [33]. Esta base de datos se encuentra actualmente disponible a través de Internet de forma totalmente libre. La base de datos contiene datos de sensores, estadísticos de aplicaciones y de voz de 21 usuarios durante más de 60 días. La tipología de los usuarios es diversa, hombres y mujeres, con edades comprendidas entre los 18 y los 70 años.

Una vez se había obtenido la base de datos, se exploró qué tan precisas son cada una de las dimensiones elegidas, resolviendo así la PI1. Para ello, se analizaron y evaluaron diversos algoritmos de detección de outliers; en el artículo se seleccionaron los tres más representativos para cada fuente. Los resultados mostraron que la voz es la fuente más sólida, con un 99.54% de AUC (Area Under Curve) y 3.40% de EER (Equal Error Rate). Desafortunadamente, la voz es también la dimensión menos presente en el uso del dispositivo y, además, los horarios en los que la voz estará disponible son imprevisibles. Tras la voz se encuentra la dimensión de los estadísticos de uso, con un 86.68% de AUC y un 20.14% de EER. Finalmente, los sensores presentan un 53.02% de AUC y un 48.83% de EER.

En este trabajo también se experimentó con la agregación de las diversas fuentes de información. Se evaluaron dos metodologías diferentes: una agregación de los vectores de

datos para construir un vector con más atributos y, por otro lado, una combinación de puntuación, donde se combinan las puntuaciones de los modelos individuales. Los resultados del segundo experimento mostraron que la combinación de puntuaciones logra una pequeña mejora con respecto a los mejores modelos individuales, y que tanto la combinación de puntuaciones como el individual son mejores que la agregación de vectores. La mejor combinación de puntuaciones resulta cuando se combinan datos de estadísticos y voz en los porcentajes 80 y 20, respectivamente. Dicha combinación consigue reducir el EER desde 2.83% a 2.59%.

Para finalizar la investigación, se decidió evaluar cómo el tamaño de la ventana de tiempo utilizada para el entrenamiento de los modelos afecta el rendimiento. Para los primeros experimentos se eligió 14 días como la cantidad de datos necesarios, a partir de la información extraída de la literatura. Las pruebas realizadas demostraron que 14 días es una buena ventana de tiempo para obtener buenos modelos, pero también que 5 días son suficientes para producir modelos con un rendimiento cercano a los entrenados con 14 días de datos recopilados. Finalmente, se evaluó el desempeño de la plataforma en una pequeña prueba de concepto. Se ha visto como el sistema, utilizando un exigente False Positive Rate (FPR) máximo del 10%, mejora notablemente sus resultados si se le añade voz. Pero el sistema necesitaba mejorar su rendimiento para poder ser llevado a cabo, dado que cuando la voz no estaba presente los ratios se veían muy mermados.

Con la intención de mejorar el rendimiento y obtener un sistema funcional, en Article 2–`IEEE_TII`, nos planteamos cuanto podría mejorar el rendimiento de la plataforma S3 del primer trabajo, utilizando ahora un enfoque supervisado. Este cambio de enfoque, nos proporcionó la posibilidad de evaluar el sistema en un entorno industrial simulado. Al ser un entorno industrial, el acceso a los datos permite entrenar los algoritmos con datos etiquetados de los usuarios, que han sido capturados en dispositivos de trabajo. A la plataforma S3 presentada en el Article 1–`Sensors`, se le aplican una serie de ligeras modificaciones para que pueda operar en dicho entorno.

El primer experimento de este trabajo se centra en analizar si el enfoque supervisado tiene (como se esperaba) mayor precisión que el no supervisado, y cuantificar en qué medida se produce la mejora. Esta experimentación resuelve la PI2. Para ello, se seleccionan algoritmos que tengan versiones para ambos enfoques. Los métodos seleccionados son KNN, RF y SVM. En este primer experimento se obtuvo una mejora significativa de más del 88% para cada uno de los diferentes tipos de datos. Por ejemplo, se pasó de un 27% a un 7.84% de EER para los datos de sensores.

Al tratarse de un entorno con similitudes a ciertos escenarios industriales, no es solo importante la precisión del sistema, sino también su robustez. Para ello se realizó un segundo experimento en el que se quería comprobar como afecta al rendimiento la aparición de nuevos trabajadores en el grupo de trabajo y así responder a la PI3. Para ello, se separó a un grupo de usuarios que representaban a usuarios desconocidos (que pueden ser nuevos trabajadores o incluso impostores). Los modelos de usuarios conocidos se entrenan únicamente con datos etiquetados del grupo de usuarios conocidos y, posteriormente, los usuarios desconocidos se evalúan como impostores. El tamaño del conjunto de usuarios desconocidos oscilaba entre 1 y 10 usuarios (lo que supone entre el 4% y el 47% del tamaño relativo del conjunto de trabajadores).

Los resultados mostraron que la degradación del rendimiento para los enfoques supervisados es más intensa cuando se usan los datos de estadísticos. La voz sigue siendo la mejor fuente de datos, dado que es la que menos se ve degradada. El False Acceptance Rate (FAR) obtenido cuando aparece solo un nuevo usuario desconocido es, como máximo (valor del intervalo de confianza superior), 1,91% para sensores, 3,49% para estadísticas

y 0,85% para voz, lo cual es muy prometedor. En cualquier caso, estos valores están por debajo del 4%, que es el tamaño relativo de los usuarios desconocidos. Los resultados de este artículo validan el sistema propuesto como un sistema candidato adecuado para la autenticación continua de los trabajadores en fábricas adheridas a la Industria 4.0.

La tercera publicación de esta tesis doctoral, Article 3–IEEE_JISA, se centra en el desarrollo e investigación de un plataforma de autenticación que preserve la privacidad de los usuarios. Un cambio importante de este trabajo respecto al enfoque anterior es pasar de una autenticación individual a una autenticación de grupo, dado que la mayoría de empresas dan permisos/roles a los usuarios según una funcionalidad que desarrollan y no siempre necesitan identificación personal. De esta forma, se preserva la identidad del usuario (y sus datos personales), dado que solo se autentica el hecho de pertenecer a un grupo de usuarios y no se revela la identidad particular. Para conseguir los objetivos de este trabajo, es necesario rediseñar la Plataforma S3 de los trabajos 1 y 2, y construir una nueva Plataforma, la CGAPP.

El primer experimento de este artículo se centra en responder a la PI4, evaluar la pérdida de precisión al aumentar la privacidad del sistema. Para ello, se desarrollan tres enfoques (niveles de privacidad) que van incrementando la privacidad, todos ellos realizados desde un enfoque de detección de outliers, metodología no supervisada para no tener que trabajar con datos etiquetados. El primer enfoque (centralizado o nivel 0) es un enfoque donde todos los datos se envían a un servidor central que genera el modelo de comportamiento del grupo. En el segundo enfoque (individual o nivel 1), cada usuario entrena un sistema propio que es enviado al sistema que los aglutina y crea el modelo del grupo. Y finalmente, el tercer enfoque (federado o nivel 2) es un sistema de entrenamiento federado, donde cada usuario entrena el sistema grupal y solo envía al servidor pesos e información del entrenamiento del modelo grupal.

Los resultados del primer experimento muestran que la plataforma CGAPP (con un nivel de privacidad 2) puede alcanzar una precisión aceptable para su funcionamiento en un entorno industrial. El mejor modelo de red neuronal de los evaluados supera el 90%, y es el mejor modelo para precisión y la métrica F1. En comparación con el enfoque centralizado (nivel de privacidad 0), donde no hay restricciones para garantizar la privacidad, la plataforma CGAPP presenta resultados inferiores, pero se acerca al enfoque centralizado; por ejemplo, un 92% frente al 96%. En cuanto al nivel de privacidad 1, donde se utiliza el enfoque individual, los resultados muestran claramente que el rendimiento del sistema se degrada significativamente.

Una vez comprobado en el primer experimento que CGAPP, con un nivel de privacidad 2, obtiene métricas similares al nivel de privacidad 0, es el momento de responder a la PI5 y evaluar la resistencia ante ataques adversariales. El primer ataque adversario estudiado son los ataques de inyección, donde un atacante utiliza el dispositivo de otro usuario (comprometido) para inyectar sus propios datos. Los resultados de este segundo experimento demuestran la diferente naturaleza de los usuarios, dado que se pueden observar comportamientos muy diferentes entre ellos. Entre las peculiaridades destaca un atacante que, independientemente del trabajador comprometido que utilice, consigue ser aceptado el 30% de las veces. Sin embargo, por mucho que aumente su ataque, no consigue aumentar su porcentaje de éxito. Otro resultado a destacar de este experimento es que ninguno de los trabajadores atacados detectarían que están siendo comprometidos, porque los atacantes no obstaculizan sus tasas de autenticación. Al contrario, incluso las mejoran. Vista la peligrosidad de este tipo de ataque, en el mismo experimento se evalúan diferentes contramedidas para paliar los ataques respondiendo a la PI6, los resultados de estas contramedidas muestran que alivian un poco los ataques, pero no consiguen contrarrestarlos

del todo.

Continuando con los ataques adversariales, en el tercer experimento de este trabajo se analizan los ataques de perturbación de datos. En estos ataques, un usuario comprometido contamina sus datos de forma que el modelo quede inutilizado, inservible. En el experimento se evalúan dos metodologías diferentes de contaminación de datos: o bien el atacante usa datos de su propia distribución, o bien usa datos de la distribución del grupo de usuarios. Se evalúan estos dos métodos, debido a que son ataques fáciles de llevar a cabo y difíciles de detectar, ya que los datos contaminados siguen la misma distribución que los legítimos. Los resultados del experimento muestran que, para el primer tipo de contaminación, el uso de su propia distribución, apenas se degradaría el sistema. Para verse afectado el sistema, deberían realizar este mismo tipo de ataque al menos cuatro usuarios al mismo tiempo y contaminar más del 50% de sus muestras. Por otro lado, al usar datos de la distribución global del grupo el sistema se degrada rápidamente; apenas es necesario un usuario y unas pocas muestras contaminadas. Este tipo de ataque sería muy peligroso, pues el sistema no rechazaría nunca la entrada a un usuario legítimo (nunca se percatarían de un mal funcionamiento), mientras que sí autenticaría a los atacantes. Ninguna de las contramedidas evaluadas palía los efectos de este ataque.

V Conclusiones y trabajo futuro

La tesis doctoral, formada por los tres artículos de investigación que la componen: Article 1–`Sensors`, Article 2–`IEEE_TII` y Article 3–`IEEE_JISA`, proporciona una visión detallada y exhaustiva de la autenticación continua en entornos industriales, abordando las preguntas de investigación planteadas en la Sección I de esta introducción y logrando cumplir con los objetivos establecidos en su Sección II.

En primer lugar, se ha demostrado que la autenticación continua es una solución altamente prometedora para mejorar la seguridad. Los resultados obtenidos en los tres estudios respaldan esta afirmación al proporcionar pruebas sólidas de que los sistemas de autenticación continua pueden ser eficaces en la protección del entorno industrial manteniendo la integridad de los datos y la privacidad de los trabajadores.

Cada uno de los tres artículos enfoca diferentes aspectos de la autenticación continua, explorando diversas fuentes de datos, como sensores, estadísticas y voz. Se ha demostrado que cada fuente de datos tiene sus propias ventajas y desafíos. Además, se ha comparado el rendimiento de los enfoques supervisados y no supervisados, y se ha validado que los enfoques supervisados mejoran significativamente la precisión de los sistemas. Al mismo tiempo se ha evaluado la generalización o robustez de estos sistemas ante nuevas incorporaciones de trabajadores. Los resultados proporcionan una base sólida para futuras investigaciones que podrían centrarse en la mejora del enfoque supervisado y la recopilación de datos etiquetados.

La privacidad de los datos de los usuarios se ha ido convirtiendo en un aspecto crítico en la sociedad de la información, en general, y en la aplicación de las técnicas de aprendizaje automático de la inteligencia artificial, en particular. Conscientes de que se convertirá en un requisito, hemos considerado este condicionante en la aplicación de la autenticación continua, abordándolo con éxito mediante el enfoque de aprendizaje federado a un sistema de autenticación grupal. La unión de la autenticación grupal junto con un esquema de entrenamiento federado protegen tanto la identidad como los datos del usuario. Por un lado, la autenticación grupal impide que sea conocida su identidad y, por otro, el entrenamiento federado garantiza que los datos se mantengan en los dispositivos personales de los usuarios, protegiendo su privacidad. Sin embargo, también se han identificado desafíos

y limitaciones, lo que sugiere que futuros trabajos podrían centrarse en mejorar aún más las consideraciones de privacidad.

La robustez de los sistemas de autenticación continua ante ataques es otra área de interés en estos estudios. Se han evaluado diferentes tipos de ataques, incluyendo ataques de inyección de datos y de perturbación de datos. Si bien los sistemas han demostrado ser robustos en el entorno, se ha identificado la necesidad de desarrollar contramedidas efectivas para garantizar la seguridad en situaciones de ataque.

En términos de implementación, se ha demostrado que los sistemas de autenticación continua son viables en entornos industriales y no afectan significativamente a la productividad de los trabajadores. Esto sugiere que los sistemas podrían ser desplegados en entornos industriales reales en un futuro cercano.

Para el trabajo futuro, se plantean diversas áreas de enfoque. En primer lugar, se buscará mejorar la precisión y la seguridad de los sistemas, incluyendo la optimización de algoritmos y la recopilación de datos etiquetados. Además, se explorarán nuevas características y fuentes de datos para una autenticación continua más sólida. La consideración de desafíos de privacidad y la mejora de las contramedidas de seguridad son áreas de interés prevalentes ahora y lo serán aún más en el futuro. También se buscará llevar estos sistemas a la producción en entornos industriales reales, lo que implica la necesidad de acuerdos y colaboraciones con empresas. En resumen, estos estudios representan un hito importante en la investigación de la autenticación continua en entornos industriales, han expandido el estado del arte y establecen una sólida base para futuras investigaciones y mejoras en esta área en constante evolución.

Bibliografía

[1] S. Almalki, P. Chatterjee, and K. Roy, "Continuous authentication using mouse clickstream data analysis," in *Security, Privacy, and Anonymity in Computation, Communication, and Storage*, G. Wang, J. Feng, M. Z. A. Bhuiyan, and R. Lu, Eds. Cham: Springer International Publishing, 2019. ISBN 978-3-030-24900-7 pp. 76–85.

[2] P. C. Alcaraz, P. Y. Zhang, P. A. Cardenas, and P. L. Zhu, "Guest editorial: Special section on security and privacy in industry 4.0," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6530–6531, 2020. DOI: 10.1109/TII.2020.2990878

[3] Z. Gao, A. Castiglione, and M. Nappi, "Guest editorial: Biometrics in industry 4.0: Open challenges and future perspectives," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 9068–9071, 2022. DOI: 10.1109/TII.2022.3197691

[4] R. Spolaor, Q. Li, M. Monaro, M. Conti, L. Gamberini, and G. Sartori, "Biometric authentication methods on smartphones: A survey," *PsychNology J.*, vol. 14, pp. 87–98, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:34009370

[5] L. Hernández-Álvarez, J. M. de Fuentes, L. González-Manzano, and L. Hernández Encinas, "Smartcampp - smartphone-based continuous authentication leveraging motion sensors with privacy preservation," *Pattern Recognition Letters*, vol. 147, pp. 189–196, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865521001434. DOI: https://doi.org/10.1016/j.patrec.2021.04.013

[6] L. Gonzalez-Manzano, J. M. D. Fuentes, and A. Ribagorda, "Leveraging user-related internet of things for continuous authentication: A survey," *ACM Comput. Surv.*, vol. 52, no. 3, Jun. 2019. [Online]. Available: https://doi.org/10.1145/3314023. DOI: 10.1145/3314023

[7] A. Acien, A. Morales, R. Vera-Rodriguez, and J. Fierrez, *Mobile Active Authentication based on Multiple Biometric and Behavioral Patterns.* Cham: Springer International Publishing, 2020, pp. 161–177. ISBN 978-3-030-39489-9. [Online]. Available: https://doi.org/10.1007/978-3-030-39489-9_9

[8] Y. Li, H. Hu, Z. Zhu, and G. Zhou, "Scanet: sensor-based continuous authentication with two-stream convolutional neural networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 16, no. 3, pp. 1–27, 2020.

[9] Y. Li, J. Luo, S. Deng, and G. Zhou, "Cnn-based continuous authentication on smartphones with conditional wasserstein generative adversarial network," *IEEE Internet of Things Journal*, vol. 9, no. 7, pp. 5447–5460, 2022. DOI: 10.1109/JIOT.2021.3108822

[10] L. Wang, M. S. Hossain, J. Pulfrey, and L. Lancor, "The effectiveness of zoom touchscreen gestures for authentication and identification and its changes over time," *Computers & Security*, vol. 111, p. 102462, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167404821002868. DOI: https://doi.org/10.1016/j.cose.2021.102462

[11] H. C. Volaka, G. Alptekin, O. E. Basar, M. Isbilen, and O. D. Incel, "Towards continuous authentication on mobile phones using deep learning models," *Procedia Computer Science*, vol. 155, pp. 177 – 184, 2019, the 16th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2019),The 14th International Conference on Future Networks and Communications (FNC-2019),The 9th International Conference on Sustainable Energy Information Technology. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S187705091930941X. DOI: https://doi.org/10.1016/j.procs.2019.08.027

[12] N. Mehdi and B. Starly, "Witness box protocol: Automatic machine identification and authentication in industry 4.0," *Computers in Industry*, vol. 123, p. 103340, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361520305741. DOI: https://doi.org/10.1016/j.compind.2020.103340

[13] I. Islam, K. M. Munim, M. N. Islam, and M. M. Karim, "A proposed secure mobile money transfer system for sme in bangladesh: An industry 4.0 perspective," in *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*, 2019, pp. 1–6. DOI: 10.1109/STI47673.2019.9068075

[14] Y. Borgianni, E. Rauch, L. Maccioni, and B. G. Mark, "User experience analysis in industry 4.0 - the use of biometric devices in engineering design and manufacturing," in *2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2018, pp. 192–196. DOI: 10.1109/IEEM.2018.8607367

[15] W. Shi, J. Yang, Yifei Jiang, Feng Yang, and Yingen Xiong, "Senguard: Passive user identification on smartphones using multiple sensors," in *2011 IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2011, pp. 141–148. DOI: 10.1109/WiMOB.2011.6085412

[16] I. Papavasileiou, Z. Qiao, C. Zhang, W. Zhang, J. Bi, and S. Han, "Gaitcode: Gait-based continuous authentication using multimodal learning and wearable sensors," *Smart Health*, p. 100162, 2020.

[17] J. Jorquera Valero, P. Sánchez Sánchez, L. Fernández Maimó, A. Huertas Celdrán, M. Arjona Fernández, S. De Los Santos Vílchez, and G. Martínez Pérez, "Improving the security and qoe in mobile devices through an intelligent and adaptive continuous authentication system," *Sensors*, vol. 18, no. 11, p. 3769, Nov 2018. [Online]. Available: http://dx.doi.org/10.3390/s18113769. DOI: 10.3390/s18113769

[18] P. M. Sánchez Sánchez, L. Fernández Maimó, A. Huertas Celdrán, and G. Martínez Pérez, "Authcode: A privacy-preserving and multi-device continuous authentication architecture based on machine and deep learning," *Computers & Security*, vol. 103,

p. 102168, 2021. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167404820304417. DOI: https://doi.org/10.1016/j.cose.2020.102168

[19] T. Feng, X. Zhao, N. DeSalvo, Z. Gao, X. Wang, and W. Shi, "Security after login: Identity change detection on smartphones using sensor fusion," in *2015 IEEE International Symposium on Technologies for Homeland Security (HST)*, 2015, pp. 1–6. DOI: 10.1109/THS.2015.7225268

[20] H. Crawford, K. Renaud, and T. Storer, "A framework for continuous, transparent mobile device authentication," *Computers & Security*, vol. 39, pp. 127 – 136, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167404813000886. DOI: https://doi.org/10.1016/j.cose.2013.05.005

[21] X. Jin and J. Han, *K-Means Clustering*. Boston, MA: Springer US, 2010, pp. 563–564. ISBN 978-0-387-30164-8. [Online]. Available: https://doi.org/10.1007/978-0-387-30164-8_425

[22] Y. Barlas, O. E. Basar, Y. Akan, M. Isbilen, G. I. Alptekin, and O. D. Incel, "Dakota: Continuous authentication with behavioral biometrics in a mobile banking application," in *2020 5th International Conference on Computer Science and Engineering (UBMK)*, 2020, pp. 1–6. DOI: 10.1109/UBMK50275.2020.9219365

[23] N. Cristianini and E. Ricci, *Support Vector Machines*. Boston, MA: Springer US, 2008, pp. 928–932. ISBN 978-0-387-30162-4. [Online]. Available: https://doi.org/10.1007/978-0-387-30162-4_415

[24] L. Breiman, "Random forest," *Machine Learning*, vol. 45, pp. 5–32, 2001. DOI: 10.1109/MSEC.2020.3039941

[25] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[26] L. Harn, "Group authentication," *IEEE Transactions on Computers*, vol. 62, no. 9, pp. 1893–1898, 2013. DOI: 10.1109/TC.2012.251

[27] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, A. Singh and J. Zhu, Eds., vol. 54. PMLR, 20–22 Apr 2017, pp. 1273–1282. [Online]. Available: https://proceedings.mlr.press/v54/mcmahan17a.html

[28] M. S. Jere, T. Farnan, and F. Koushanfar, "A taxonomy of attacks on federated learning," *IEEE Security & Privacy*, vol. 19, no. 2, pp. 20–28, 2021. DOI: 10.1109/MSEC.2020.3039941

[29] X. Cao and N. Gong, "Mpaf: Model poisoning attacks to federated learning based on fake clients," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Los Alamitos, CA, USA: IEEE Computer Society, jun 2022, pp. 3395–3403. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/CVPRW56347.2022.00383. DOI: 10.1109/CVPRW56347.2022.00383

[30] J. Fan, Q. Yan, M. Li, G. Qu, and Y. Xiao, "A survey on data poisoning attacks and defenses," in *2022 7th IEEE International Conference on Data Science in Cyberspace (DSC)*, 2022, pp. 48–55. DOI: 10.1109/DSC55868.2022.00014

[31] H. T. Reda, A. Anwar, and A. Mahmood, "Comprehensive survey and taxonomies of false data injection attacks in smart grids: attack models, targets, and impacts," *Renewable and Sustainable Energy Reviews*, vol. 163, p. 112423, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032122003306. DOI: https://doi.org/10.1016/j.rser.2022.112423

[32] S. A. Alsuhibany, "A survey on adversarial perturbations and attacks on captchas," *Applied Sciences*, vol. 13, no. 7, 2023. [Online]. Available: https://www.mdpi.com/2076-3417/13/7/4602. DOI: 10.3390/app13074602

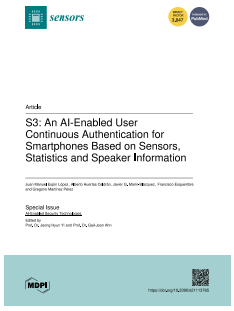
[33] J. M. E. López, A. H. Celdrán, J. G. Marín-Blázquez, F. E. Martínez, and G. M. Pérez, "S3 Dataset," 4 2021. [Online]. Available: https://figshare.com/articles/dataset/S3Dataset_zip/14410229. DOI: 10.6084/m9.figshare.14410229.v2

Publicaciones que componen

la Tesis Doctoral (PhD Tesis)

1

S3: An AI-Enabled User Continuous Authentication for Smartphones Based on Sensors, Statistics and Speaker Information



Title:	S3: An AI-Enabled User Continuous Authentication for Smartphones Based on Sensors, Statistics and Speaker Information
Authors:	Juan Manuel Espín López[1], Alberto Huertas Celdrán[2], Javier G. Marín-Blázquez[1], Francisco Esquembre[3], Gregorio Martínez Pérez[1]
Filiación:	1- Departamento de Informática e Ingeniería de las Comunicaciones (DIIC), Universida de Murcia, 2-Communication Systems Group (CSG), Departamento Informática (IfI), Universidad de Zürich, 3-Departamento de Matemáticas, Universidad de Murcia.
Journal:	Sensors
JIF:	3.847 Q2 (2021)
Publisher:	MDPI
Volume:	21
Number:	11
Pages:	3765
Year:	2021
Month:	May
DOI:	10.3390/s21113765
Status:	Published

Abstract

Continuous authentication systems have been proposed as a promising solution to authenticate users in smartphones in a non-intrusive way. However, current systems have important weaknesses related to the amount of data or time needed to build precise user profiles, together with high rates of false alerts. Voice is a powerful dimension for identifying subjects but its suitability and importance have not been deeply analyzed regarding its inclusion in continuous authentication systems. This work presents the S3 platform, an artificial intelligence-enabled continuous authentication system that combines data from sensors,

applications statistics and voice to authenticate users in smartphones. Experiments have tested the relevance of each kind of data, explored different strategies to combine them, and determined how many days of training are needed to obtain good enough profiles. Results showed that voice is much more relevant than sensors and applications statistics when building a precise authenticating system, and the combination of individual models was the best strategy. Finally, the S3 platform reached a good performance with only five days of use available for training the users' profiles. As an additional contribution, a dataset with 21 volunteers interacting freely with their smartphones for more than sixty days has been created and made available to the community.

Keywords

Continuous Authentication · Smartphone · Sensors · Applications Usage · Speaker Recognition · Artificial Intelligence

2

A Supervised ML Biometric Continuous Authentication System for Industry 4.0

Title:	A Supervised ML Biometric Continuous Authentication System for Industry 4.0
Authors:	Juan Manuel Espín López[1], Alberto Huertas Celdrán[2], Francisco Esquembre[3], Gregorio Martínez Pérez[1], Javier. G. Marín-Blázquez[1]
Filiación:	1- Departamento de Informática e Ingeniería de las Comunicaciones (DIIC), Universida de Murcia, 2-Communication Systems Group (CSG), Departamento Informática (IfI), Universidad de Zürich, 3-Departamento de Matemáticas, Universidad de Murcia.
Journal:	IEEE Transaction on Industrial Informatics
JIF:	12.300 D1 (2022)
Publisher:	IEEE
Volume:	18
Number:	12
Pages:	9132-9140
Year:	2022
Month:	Dic
DOI:	10.1109/TII.2022.3171321
Status:	Published

Abstract

Continuous authentication (CA) is a promising approach to authenticate workers and avoid security breaches in the industry, especially in Industry 4.0, where most interaction between workers and devices takes place. However, introducing CA in industries raises the following unsolved questions regarding machine learning (ML) models: its precision and performance; its robustness; and the issue about if or when to retrain the models. To answer these questions, this article explores these issues with a proposed supervised versus nonsupervised ML-based CA system that uses sensors, applications statistics, or speaker data collected by the operator's devices. Experiments show supervised models with equal error rates of 7.28% using sensors data, 9.29% with statistics, and 0.31% with voice, a sig-

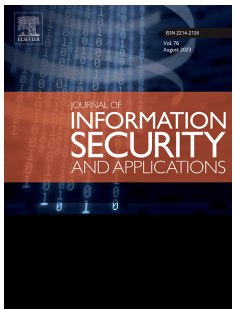
nificant improvement of 71.97, 62.14, and 97.08%, respectively, over unsupervised models. Voice is the most robust dimension when adding new workers, with less than 2% of false acceptance rate even if workforce size is doubled.

Keywords

Cybersecurity · Safety · Neuronal cyberattacks · Convolutional neural networks · Brain-computer interfaces Applications Usage · Continuous Authentication (CA) · Industry 4.0 · Machine Learning (ML)/Deep Learning (DL) · Sensors · Speaker Recognition

3

CGAPP: A Continuous Group Authentication Privacy-Preserving Platform for Industrial Scene



Title:	CGAPP: A continuous group authentication privacy-preserving platform for industrial scene
Authors:	Juan Manuel Espín López[1], Alberto Huertas Celdrán[2], Francisco Esquembre[3], Gregorio Martínez Pérez[1], Javier. G. Marín-Blázquez[1]
Filiación:	1- Departamento de Informática e Ingeniería de las Comunicaciones (DIIC), Universida de Murcia, 2-Communication Systems Group (CSG), Departamento Informática (IfI), Universidad de Zürich, 3-Departamento de Matemáticas, Universidad de Murcia.
Journal:	Journal of Information Security and Applications
JIF:	5.6 Q2 (2022)
Publisher:	Elsevier Ltd.
Volume:	78
Pages:	103622
Year:	2023
Month:	Nov
DOI:	10.1016/j.jisa.2023.103622
Status:	Published

Abstract

In Industry 4.0, security begins with the workers' authentication, which can be done individually or in groups. Recently, group authentication is gaining momentum, allowing users to authenticate as group members without the need to specify the particular individual. Continuous authentication and federated learning are promising techniques that might help group authentication by providing privacy, by its own design, and extra security compared to traditional methods based on passwords, tokens, or biometrics. However, these techniques have not previously been combined or evaluated for authenticating workers in Industry 4.0. Thus, this paper proposes a novel continuous group authentication privacy-preserving (CGAPP)platform that is suitable for the industry. The CGAPP platform incorporates statistical data from workers' smartphones and employs federated learning-based outlier detection for group worker authentication while ensuring the privacy

of personal data vectors. A series of experiments were performed to measure the framework's suitability and address the following research questions: (i) What is the cost of using FL compared to full data access in industrial scenarios? (ii) How robust is federated learning against adversarial attacks, specifically, how much malicious data is required to deceive the model? and (iii) How much noise is required to disrupt the authentication system? The results demonstrate the effectiveness of the CGAPP platform in the industry since it provides factory safety while preserving privacy. This platform achieves an accuracy of 92%, comparable to the 96% obtained by traditional approaches in the literature that do not address privacy concerns. The platform's robustness is tested against attacks in the second and third experiments, and various countermeasures are evaluated. While the CGAPP platform exhibits certain vulnerabilities to data injection attacks, straightforward countermeasures can alleviate them. Nevertheless, the system's performance experiences a notable impact in the event of a data perturbation attack, and the countermeasures investigated are ineffective in addressing this issue.

Keywords

Continuous authentication · Group authentication · Federated learning · Adversarial attack · Industry 4.0