CAPÍTULO 1. INTRODUCCIÓN

1.1. Contexto de la investigación

La necesidad de proteger los activos y recursos frente a accesos indebidos de usuario no autorizados siempre ha existido. Del mismo modo que en un entorno físico como un aeropuerto, se necesita controlar que personas están autorizadas a salir o entrar de un país, en un sistema de información existe esta misma necesidad de controlar el acceso a sus recursos.

En el ámbito digital, la disciplina que se encarga de gestionar las diferentes identidades de los usuarios y la información relacionada con las mismas, con el objetivo de controlar el acceso a los diferentes recursos, servicios o aplicaciones, se denomina gestión de identidades y accesos [1]. Esta disciplina trata de garantizar la seguridad de los usuarios y de los recursos, servicios o aplicaciones, ya sea de manera individual o aislada, o al mismo tiempo en multitud de ellos alojados en diversos dominios.

Los primeros sistemas de información estaban pensados para ser utilizados por un único usuario. De esta forma, el usuario encargado de utilizar el sistema debía autenticarse con el objetivo de validar sus credenciales y, por lo tanto, pasar a comprobar los permisos y privilegios disponibles para interactuar con dicho sistema. En otras palabras, dicho usuario debía autenticarse en el sistema con el objetivo de poder autorizar las acciones que quería realizar en dicho sistema de información. Con el paso del tiempo, se empezaron a implantar los sistemas multiusuario, en los que diversos usuarios interactúan simultáneamente con el mismo sistema y pueden estar autorizados a realizar diferentes operaciones dependiendo de su nivel de privilegios. De esta forma, los recursos computacionales han de repartirse en tiempo y espacio entre todos los usuarios de forma totalmente transparente para ellos. Así, los sistemas de gestión de identidades tuvieron que evolucionar para dar respuesta a los requisitos de los nuevos casos de uso.

Desde la llegada de Internet la gestión de identidades y accesos se fue adaptando para considerar que un número elevado de usuarios pueden estar solicitando acceso a un mismo servicio, aplicación o recurso alojados en servidores externos, heterogéneos y completamente distribuidos, haciendo que su gestión sea más complicada. Esto cambió el paradigma, pues la gestión de identidades se estaba volviendo demasiado compleja debido al gran número de identidades digitales que cada usuario tenía que manejar de forma independiente, ya que para cada servicio o aplicación se tenía que generar una nueva identidad (con su cuenta y credenciales asociadas).

Los últimos mecanismos desarrollados e implantados para solventar la gestión de identidades y accesos se basan en el concepto de federación. Estos mecanismos se basan en estándares y protocolos que permiten que un mismo usuario pueda utilizar una misma identidad digital para realizar cualquier flujo de Identificación, Autenticación, Autorización y Auditoria (IAAA) a lo largo de multitud de servicios, recursos o aplicaciones, alojados en diversos dominios, utilizando para ello el mismo proveedor de identidades de confianza. Esto se consigue porque los proveedores de servicios y aplicaciones delegan los procesos de IAAA en el proveedor de identidades. Sin embargo, esto no quiere decir que los recursos, servicios o aplicaciones deleguen totalmente la seguridad de los usuarios en manos de los proveedores de identidades y, por lo tanto, también han de asumir su responsabilidad a la hora de proteger a los usuarios.

¿De qué se debe proteger a los usuarios en este contexto? Principalmente de sufrir una suplantación de identidad, es decir, de situaciones en las que un tercero malicioso pueda actuar en su nombre (suplantando su identidad) para actuar con un recurso, servicio o aplicación. Este tipo de suplantaciones suelen producirse cuando las credenciales del usuario se ven comprometidas (por ejemplo, por un ataque de *phishing* o por una brecha de datos) o cuando su sesión se secuestra (por ejemplo, por un ataque de tipo *TCP hijacking* o por uno de *Cross Site Scripting* o *Cross Site Request Foegery*).

Para detectar o incluso evitar este tipo de ataques, esta tesis doctoral pretende explorar la aplicación de técnicas de análisis de comportamientos. Los fundamentos del análisis de comportamientos (en inglés, *behavioral analysis*) fueron introducidos por primera vez en 1953 [2]. En aquel entonces esta rama estaba muy ligada al campo de la psicología y se basaba principalmente en el análisis del comportamiento humano con el objetivo de poder aportar un beneficio concreto. Sus fundamentos principales fueron determinados posteriormente en 1960 [3].

Hoy en día, el análisis de comportamiento se denomina analítica del comportamiento (en inglés, *behavioral analytics*). La analítica del comportamiento se define como la rama científica que se centra en modelar el comportamiento de una entidad (usuario, animales, sensores, etc) para poder analizar y entender las interacciones que realiza en un sistema de información, con el objetivo de obtener un beneficio de negocio [4]. De esta forma, se hace uso del aprendizaje máquina para modelar y predecir los comportamientos pasados, presentes y futuros de las entidades.

Las técnicas del análisis de comportamiento se pueden categorizar en función del tipo de entidad que se considera como objeto de estudio. De esta forma, existe el análisis de comportamiento animal [5], análisis de comportamiento

de sensores [6] y el análisis de comportamiento humano [7]. El análisis de comportamiento animal se basa en utilizar sensores acoplados a animales para comprender, monitorizar y predecir determinadas actividades de interés, por ejemplo, en una granja animal se utilizan collares inteligentes para monitorizar la actividad del ganado y predecir comportamientos que puedan ser indicadores de que el ganado está enfermo [8]. El análisis de comportamiento de sensores se basa en modelar las medidas captadas por ciertos sensores con el objetivo de detectar anomalías [9] para diversos fines como, por ejemplo, realizar procesos de autenticación [10]. Hoy en día, estas ramas del análisis de comportamiento están evolucionando constantemente gracias a la aparición del *Internet of Things* (IoT) y al desarrollo y aumento del número de sensores de bajo coste que se pueden acoplar en multitud de dispositivos del entorno físico de forma sencilla.

Por otro lado, el análisis de comportamiento humano se basa en modelar el comportamiento de personas o usuarios, tanto a nivel individual [11], como en colectivos o grupos sociales [12]. Cuando el dominio de aplicación es un sistema de información específico, el área científica se denomina análisis de comportamiento de usuarios.

Y es en este ámbito en el que se centra la presente tesis doctoral, ya que las técnicas de análisis de comportamiento deberían proporcionar una gran ventaja frente a los modelos convencionales a la hora de detectar intrusiones, ataques o fraudes.

1.2. Hipótesis de partida

Las especificaciones de gestión de identidades y accesos federada están siendo adoptadas muy rápidamente y a gran escala para solventar los procesos de IAAA. Algunos ejemplos claros se pueden ver en productos como Google Connect [13], Facebook Connect [14] o en el sector bancario con el desarrollo de *Financial-grade API* (FAPI) [15]. Sin embargo, estas especificaciones y sus implementaciones presentan vulnerabilidades de seguridad que se pueden explotar causando daños irreversibles en los sistemas de información [16]. Las principales líneas de investigación y soluciones planteadas hasta el momento para mitigar todas estas amenazas suelen estar orientadas al enriquecimiento de las peticiones y/o *tokens* empleados en los estándares, a realizar una mejora en la gestión de las sesiones de usuario, a la mejora de las *Application Programming Interface*s (APIs) y *Software Development Kit*s (SDKs) ofrecidas, o al uso de criptografía a diferentes niveles [17].

Los modelos de análisis de comportamiento se presentan como una solución efectiva y viable para solventar problemas relacionados con los procesos de

IAAA [18]. Estas soluciones, no sólo permiten automatizar ciertos procesos, sino que también se posicionan como una de las soluciones más eficaces a la hora de detectar brechas de seguridad. Aun siendo una buena solución, todavía existe margen de mejora en los modelos existentes, especialmente en la combinación de información de múltiples fuentes de información heterogéneas.

Teniendo en cuenta estas dos líneas, surge como principal motivación de esta investigación tratar de integrar los modelos de análisis de comportamientos dentro de los principales estándares de gestión de identidades federada. Esto ayudaría a aumentar los niveles de seguridad proporcionados en dichos estándares, además de evitar que los recursos, servicios o aplicaciones deleguen totalmente la seguridad en los proveedores de identidades. Por otro lado, se pueden mejorar los modelos de análisis de comportamientos existentes, con el objetivo de que sean más eficaces y puedan escalar de forma sencilla para considerar simultáneamente multitud de fuentes de información. Por todo esto, la hipótesis de partida de la presente tesis doctoral es:

Es posible mejorar los niveles de seguridad proporcionados actualmente por los principales esquemas de gestión de identidades federada integrando métodos de análisis de comportamientos en los flujos de información. Además, es posible mejorar la eficacia de los métodos de análisis de comportamiento actuales generando modelos que puedan combinar información de múltiples fuentes de datos heterogéneas de forma efectiva.

1.3. Objetivos

Los objetivos que la presente tesis doctoral pretende conseguir surgen de su hipótesis de partida. Estos objetivos generales se resumen en:

- Diseñar, implementar y evaluar un flujo de trabajo que defina y marque las pautas a seguir para integrar métodos de análisis de comportamientos en los flujos de información de los esquemas de gestión de identidades federada.

- Diseñar, implementar y evaluar un método de análisis de comportamientos que permita combinar información de fuentes de datos heterogéneas, mejorando la eficacia de los modelos del estado del arte.

Para conseguir el primer objetivo general se han definido los siguientes objetivos específicos:

1. Definir y esquematizar las principales tareas que ha de seguir una aplicación o servicio para integrar una solución de análisis de comportamientos en los estándares federados.

2. Definir los procesos de obtención y generación de huellas digitales que mejor representen el comportamiento de los usuarios de un sistema de información.

3. Definir y categorizar los posibles modelos de aprendizaje máquina y las métricas de evaluación de interés en el ámbito del análisis de comportamiento.

4. Analizar los posibles métodos de acción para integrar la información recogida de los modelos de análisis de comportamientos en los flujos de información de los estándares federados.

5. Validar y evaluar el flujo de trabajo propuesto en un entorno real.

Del mismo modo, para el segundo objetivo general, se definen los siguientes objetivos específicos:

1. Generar un conjunto de datos nuevo para aliviar la escasez y falta de datos en el ámbito del análisis de comportamiento.

2. Definir e implementar una técnica para representar la información de comportamientos que permita combinarla a nivel de características.

3. Definir e implementar una técnica que permita mejorar la eficiencia de los modelos de análisis de comportamientos.

4. Proponer un modelo de riesgos que, considerando las técnicas anteriores, sea capaz de detectar anomalías de comportamiento de forma efectiva.

5. Validar y evaluar el método basado en estas técnicas y modelo en un entorno real.

1.4. Metodología

La metodología utilizada con el fin de cumplir los objetivos planteados y poder demostrar la hipótesis de partida se define como sigue:

- Análisis y comprensión los flujos de información definidos por los esquemas de gestión de identidades federados.

- Análisis de los trabajos previos en el ámbito del análisis de comportamiento con el objetivo de encontrar limitaciones en los mismos cuando se aplican al dominio de interés en esta tesis.

- Análisis de los tipos de huellas digitales que se utilizan en la literatura para seleccionar los atributos más representativos y singulares.

- Análisis de las peculiaridades de los datos de dinámicas de comportamientos para seleccionar las métricas de evaluación que mejor se adapten y caractericen la problemática definida.

- Análisis y comprensión de las técnicas de combinación de la información en el ámbito del aprendizaje máquina.

- Creación de un entorno de trabajo real. Este entorno ha de implementar un sistema de gestión de identidades federado así como, un servicio o aplicación que haga uso de él.

- Recolección de los datos que contienen dinámicas de comportamientos.

- Evaluación y validación del flujo de trabajo propuesto utilizando el entorno de trabajo. Repercusión y evaluación de la seguridad añadida que se proporciona, así como de las posibles latencias añadidas que supone su implantación.

- Evaluación y validación del método de combinación de la información utilizando los datos recolectados.

- Comparativa del método propuesto con otros trabajos y algoritmos del estado del arte.

1.5. Estructura del documento

El presente documento de tesis doctoral se compone de un total de seis capítulos. En este primer capítulo se introducen los primeros conceptos relacionados con las temáticas relacionadas con la tesis, es decir, la gestión de identidades y accesos y el análisis de comportamientos. Además, se ha definido la hipótesis de partida, los objetivos tanto generales como específicos y la metodología a seguir para poder demostrar la hipótesis de partida. De aquí en adelante, los cinco capítulos restantes se resumen en:

- En el **Capítulo 2** se exponen y analizan los trabajos del estado del arte relacionados con las dos líneas de investigación de interés. En primer lugar,

se abordan los conceptos fundamentales del campo del control de accesos y se recorre la evolución de los sistemas de gestión de identidades hasta llegar a los esquemas federados. Posteriormente se profundiza en el análisis de comportamiento, partiendo de los principales dominios de aplicación hasta llegar al ámbito de la ciberseguridad. A continuación, se presentan los trabajos relacionados con el análisis de comportamientos para solventar problemas específicos de control de accesos, mayoritariamente problemas de autenticación utilizando fuentes de datos como el teclado y el ratón. Finalmente, se detallan las limitaciones encontradas en la literatura.

- En el **Capítulo 3** se propone un flujo de trabajo que permite integrar los métodos de análisis de comportamiento en los principales estándares de gestión de identidades federados. En primer lugar, se exponen los casos de uso en los que el flujo de trabajo puede mejorar los niveles de seguridad proporcionados actualmente. Posteriormente, se detallan todas las tareas que componen dicho flujo de trabajo y que marcan las directrices, consideraciones y recomendaciones que se deben seguir para una correcta implantación.

- En el **Capítulo 4** se propone un método para combinar información de múltiples fuentes de datos heterogéneas en el ámbito del análisis de comportamiento. Este método permite detectar anomalías de comportamiento y, por lo tanto, detectar brechas de seguridad y mejorar los sistemas de control de accesos. De esta forma, se definen y detallan todos los algoritmos y técnicas de aprendizaje máquina empleadas con el objetivo de generar un modelo robusto, con una eficacia elevada y capaz de generalizar de forma correcta.

- En el **Capítulo 5** se exponen los diferentes experimentos llevados a cabo para poder evaluar y validar las propuestas definidas en los Capítulos 3 y 4. En primer lugar, se detalla la creación de un entorno de trabajo. Este entorno permite la creación de un conjunto de datos que contiene dinámicas de comportamiento. Posteriormente, se procede a evaluar por separado tanto el flujo de trabajo, como el método de combinación de la información.

- En el **Capítulo 6** se exponen las conclusiones extraídas en la realización del presente trabajo. De esta forma, se analizan las principales lecciones aprendidas, el cumplimiento de los objetivos planteados y la validación de la hipótesis de partida. Además, se proponen futuras líneas de investigación relacionadas.

Por último, en el **apéndice A** se encuentra el glosario de acrónimos utilizados en el presente documento de tesis doctoral.