UNIVERSIDAD NACIONAL DE LA PLATA

FACULTAD DE INFORMÁTICA

TESINA DE LICENCIATURA

Título: Modelos computacionales y metodologías utilizadas en la detección del fraude de las tarjetas de crédito
Autores: Andrea Verónica Porco
Directores: Bertone Rodolfo, Thomas Pablo
Carrera: Licenciatura en sistemas

Resumen

De la mano de la evolución del comercio electrónico ha devenido el incremento en el uso de las tarjetas de crédito (TC) para realizar compras habituales en internet, y como consecuencia, se ha dado a lugar al incremento del fraude sobre las mismas, llamado fraude online.

La implementación de sistemas de detección de fraude en la administración de TC, tanto para el banco como para el usuario, tomó un lugar de privilegio, en pos de minimizar sus pérdidas. Muchas técnicas modernas basadas en inteligencia artificial, minería de datos, lógica difusa, seguridad criptográfica, identificación por datos biométricos, por GPS, por reconocimiento de rostro, por interfaces gestuales, etc., han evolucionado en esta temática, a través de la identificación del usuario, que ha ganado acceso a la sesión de aplicación del usuario titular.

Este trabajo presenta las líneas de investigación que se fueron abriendo espacio en la resolución del problema, los modelos computacionales más sobresalientes que avalan estos lineamientos, y las metodologías aplicadas a la fecha que han demostrado, serán las candidatas a resolver el problema del FTC.

Palabras Claves

Comercio electrónico, fraude de las tarjetas de crédito (FTC), detección del FTC (DFTC), modelos de DFTC (MODF), metodologías de DFTC (MEDF), fraude online, identificación del usuario, autenticación del usuario, comportamiento del usuario, información contextual.

Trabajos Realizados

Investigación del FTC y de las soluciones planteadas en la DFTC. Estudio de la problemática, razones de su incremento, análisis del impacto de la tecnología, determinación de los modelos computacionales y las metodologías que se han ido utilizando progresivamente a lo largo de los años. Comparación de soluciones, abstracción y clasificación de líneas estudiadas. Evaluación de fallas y patrones comunes. Determinación de líneas candidatas a la resolución del FTC a futuro.

Conclusiones

Los sistemas computacionales desconocen las características del usuario fraudulento.

Los modelos propuestos en la DFTC han fallado.

Se considera a la línea de investigación relacionada con la extracción de información del contexto, como la línea candidata, la que tuvo un fuerte impacto, un rápido crecimiento, y un desarrollo constante. Por todo lo analizado, se puede afirmar que esta línea, será factible se destaque a futuro en la DFTC, entre las líneas estudiadas en este trabajo.

Trabajos Futuros

Plantear un perfil de usuario fraudulento de acuerdo a sus características corporales, en comparativa al usuario titular. Proponer modelos de DFTC aplicados. Para esto, utilizar las características a mejorar en modelos anteriores, planteadas en este trabajo, con selección de al menos dos metodologías existentes en la línea candidata resultante. Realizar pruebas con varios usuarios. Analizar resultados obtenidos y posible impacto en la DFTC.

Fecha de la presentación: Marzo 2015

Índice general

Tabla de Contenidos

Tabla de figuras

Agradecimientos

Agradezco al pampa Bertone especialmente, a quien admiro profundamente, por su paciencia, su ayuda y apoyo constante para la culminación de este trabajo.

A mi papá y a mis abuelos, que son mi fuente de inspiración para toda la vida.

A mi amiga, Marta Hackbart, por alentarme en todos mis emprendimientos y estar siempre a mi lado.

A mi mamá y a mis hermanos, Micaela, Marcelo y Alejandra, que son la base de mi felicidad y a quienes amo por siempre.

A Cecilia Onaha, por sus sabios consejos y por enseñarme a ver la vida de otra manera.

A mis compañeros de la facultad y de ruta, porque me animan y alientan a seguir adelante.

Capítulo 1

Introducción

En este capítulo se describen la motivación y objetivos de la tesina

Motivación

La motivación principal de la tesina es mostrar que el marcado crecimiento del problema del FTC, ha creado un gran interés en diferentes áreas de la informática. Las áreas que han tratado este problema, lo han hecho bajo líneas de investigación específicas, definidas a través de modelos computacionales y metodologías de soporte. Estas líneas pueden incluso componerse y/o complementarse entre sí, y se diferencian con características bien marcadas, como:

- El costo de implementación de sistemas
- La accesibilidad que ofrecen
- La probabilidad de falso fraude o reporte de fallos
- La flexibilidad con que se manejan
- El uso de recursos, tanto hardware como software
- La extensibilidad, en especial de los modelos planteados
- El mantenimiento de los sistemas inmersos en determinadas líneas

Se puede observar que las líneas son netamente distintas, pero que pueden colaborar para la resolución del problema de fraude, algunas de ellas con más éxito que otras.

Objetivos de la tesina

- Identificar las líneas de investigación actuales en la DFTC.

- Mostrar el surgimiento de la problemática, explicar las razones de su incremento en los últimos años, analizar el impacto de la tecnología, determinar los modelos computacionales y las metodologías que se han ido utilizando progresivamente a lo largo de los años

- Obtener diferencias sustanciales y de impacto entre las líneas de investigación extraídas.

- Plantear las ventajas y desventajas de los modelos y las metodologías utilizados.

- Estimar a futuro la línea de investigación que será factible se destaque entre las líneas estudiadas.

Capítulo 2

El fraude de las tarjetas de crédito (FTC)

En este capítulo se explicará la definición del FTC y las etapas de la DFTC

2.1. Definición del FTC

En [MAN2008] se define un fraude a la identidad, como una explotación no autorizada de información credencial, a través del uso de una falsa identidad. Puede extenderse esta definición para hallar la definición del FTC, y decir que es la explotación no autorizada de la información credencial, en particular, aquellos datos que comprometen la TC del usuario titular, a través del uso de una falsa identidad. El problema principal a resolver, es la identificación del usuario fraudulento, para que este no gane acceso a la sesión en la aplicación usuario, donde se encuentran los datos credenciales.

2.2. Etapas de la DFTC

La globalización y el incrementado uso de Internet para compras online, ha resultado en una considerable proliferación de transacciones de TC a través del mundo. El rápido crecimiento en el número de estas transacciones, ha llevado a un incremento sustancial en actividades fraudulentas **[EDW2011]**. Por ello, el FTC es considerado un término de rango amplio, y los delincuentes que cometen este tipo de fraude, emplean un gran número de técnicas.

Lo más sobresaliente de los últimos años es el fraude online. Para combatirlo, es importante entender los mecanismos de identificación disponibles, así como las etapas en las que puede realizarse la detección.

Las etapas son las siguientes:

Etapa 1. Como primera medida, se puede detectar un intruso por la actividad que este realiza sobre la red, traspasando la seguridad establecida en la misma.

Esta etapa, apunta en gran medida a denegar el acceso del usuario fraudulento, previo a la autenticación del usuario en el sistema.

Etapa 2. Como segunda medida, se puede detectar al usuario fraudulento durante el ingreso al sistema. Esta etapa se realiza con el control de nombre de usuario y contraseña, pero podrán exigirse otras medidas complementarias, tales como, pedir información de contexto (foto de perfil, grabación de voz, ubicación del GPS, etc.) y demás requisitos establecidos por la aplicación que desea autenticar al usuario.

Etapa 3. Por último, se puede detectar al usuario mientras opera el sistema, de modo que ya ha traspasado las etapas 1 y 2, por lo que el usuario fraudulento, ya ha ganado el acceso a la aplicación del usuario titular.

La detección en esta etapa, suele realizarse tomando información del historial del usuario, tal como, movimientos de cuenta habituales, montos habituales de extracción y depósito, fechas clave de transacciones, foto del perfil anterior en comparación a foto de contexto actual, contactos a la fecha, ubicación habitual de operación, entre otros.

El tipo de identificación del usuario que puede realizarse en cada etapa, se evalúa en los modelos y metodologías de DFTC. Esto se profundiza en capítulos 4 y 5.

Capítulo 3

Fases de impacto en la DFTC

En este capítulo se explicarán las fases evolutivas y de impacto directo sobre el problema del FTC y su detección.

La evolución hacia diferentes líneas de investigación en la DFTC se presentaron progresivamente y, para evaluar estos avances, es posible definir distintas fases. En cada fase, se evalúan características que afectan la resolución de este problema.

Las fases son las siguientes:

3.1. Fase 1. Ampliación de perfiles de usuario fraudulento

En esta fase, puede observarse variados tipos de dispositivos móviles y su extensión mundial, a un costo accesible y abarcando generaciones que antes no solían hacer uso de tecnología de última generación. Esto revolucionó el uso de aplicaciones desde cualquier ubicación, y dio lugar a nuevos y variados tipos de usuario, lo que produjo un gran impacto sobre la definición de perfiles de usuarios fraudulentos.

3.2. Fase 2. El paso de la TC a la billetera electrónica.

Debido a que se realizó un cambio en el modo de operación, en compras de bienes y servicios, y en el manejo de cuentas bancarias, la detección de usuarios fraudulentos también se ve afectada. Las actividades que antes se realizaban con una TC, para el 2015, se espera que puedan realizarse sólo con el uso del dispositivo móvil, incluso cuando este se encuentre apagado. Esta modalidad de pago, asociada a una o varias cuentas bancarias, y/o cuentas establecidas con prestadores de servicios, se manifiesta en proyectos con dispositivos de tipo Near Field Communication (NFC). En estos dispositivos, el decremento del crédito que posee un usuario, se realiza por contacto cercano contra otra aplicación. Estos proyectos incluyen el uso de dispositivos NFC sobre máquinas expendedoras, de boletos en el sistema de transporte, de

alimentos y bebidas en la vía pública, de ticket en playas de estacionamiento, de compra en supermercados y centros comerciales, etc.

(Ver más en Anexo A).

Esta tendencia recae en una nueva modalidad de billetera electrónica, como sucede con algunas aplicaciones, como la aplicación de Google, llamada E-wallet, y funciona exactamente como una TC, tanto por el modo de pago, como por su modo de operación. Por este cambio en las modalidades de uso de la TC, es que la DFTC, se va incorporando a las funciones de seguridad de los dispositivos, y sobre algunas aplicaciones móviles.

Esta modalidad también, en ocasiones traspasa las etapas 1 y 2 (ver capítulo 2), debido a que, muchas de estas aplicaciones, habilitan el decremento del crédito del dispositivo con tan sólo el contacto cercano contra otra aplicación, que maneja el mismo sistema, sin autorizar al usuario previamente.

3.3. Fase 3. Se suman algunos problemas de seguridad.

La evolución en la tecnología introduce naturalmente un incremento en la inseguridad sobre la comunicación, que se debe contrarrestar, para ofrecer al usuario un nivel de seguridad aceptable en las operaciones que este realiza diariamente.

En el intento de ofrecer seguridad en las etapas 1 y 2 (ver capítulo 2) de la DFTC, se observan varias fallas, por las que los intrusos siguen accediendo a la etapa 3.

Estos son algunos problemas de seguridad:

3.3.1. El problema del phishing

El phishing es una técnica de ingeniería social, utilizada por los delincuentes para obtener información confidencial como nombres de usuario, contraseñas y/o datos financieros y bancarios, como detalles de tarjetas de crédito, haciéndose pasar por una comunicación confiable y legítima.

El escenario de phishing, generalmente está asociado con la capacidad de duplicar una página web, para hacer creer al visitante que se encuentra en el sitio web original, en lugar del falso. El engaño suele llevarse a cabo a través del correo electrónico y, a menudo estos correos contienen enlaces a un sitio web falso con una apariencia casi idéntica a un sitio legítimo. Una vez en el sitio falso, los usuarios incautos son engañados para que ingresen sus datos confidenciales, lo que le proporciona a los delincuentes un amplio margen para realizar estafas y fraudes con la información obtenida.

La principal manera de llevar adelante el engaño es a través del envío de spam (correo no deseado) e invitando al usuario a acceder a la página señuelo. A menudo, estos correos llegan a la bandeja de entrada disfrazados como procedentes de departamentos de recursos humanos o tecnología, o de áreas comerciales relacionadas a transacciones financieras.

(Ver más en Anexo B)

3.3.2. El problema del MITM (man in the middle)

Cuando se emite una información, se supone que debe llegar sólo a un receptor que se encuentra predeterminado, pero alguien con conocimiento de los protocolos de internet, puede desviar esa información hacia otra terminal, alterarla, y luego re-enviarla al receptor o receptores finales. El ataque del MITM, intercepta una comunicación entre dos sistemas. Por ejemplo, en una transacción bancaria, la intercepción es, entre el cliente y el servidor del banco.

Utilizando técnicas diferentes, el atacante divide la conexión original en 2 nuevas conexiones, una entre el cliente y el atacante, y la otra, entre el atacante y el servidor. Una vez que la conexión es interceptada, el atacante es capaz de leer y modificar los datos de la transacción bancaria, concretando el fraude.

3.3.3. Los troyanos bancarios

Se denomina troyano o caballo de Troya, a un software malicioso, que se presenta al usuario como un programa aparentemente legítimo e inofensivo, pero al ejecutarlo, le brinda a un atacante acceso remoto al equipo infectado. El término troyano proviene de la historia del caballo de Troya mencionado en la Odisea de Homero.

Los troyanos pueden realizar diferentes tareas, pero, en la mayoría de los casos crean una puerta trasera, que permite la administración remota a un usuario no autorizado.

Un troyano no es de por sí, un virus informático, aún cuando teóricamente pueda ser distribuido y funcionar como tal. La diferencia fundamental entre un troyano y un virus, consiste en su finalidad. Para que un programa sea un "troyano" sólo tiene que acceder y controlar la máquina anfitriona sin ser advertido, normalmente bajo una apariencia inocua. Al contrario que un virus, que es un huésped destructivo, el troyano no necesariamente provoca daños porque no es su objetivo.

El troyano bancario tiene por finalidad el robo de información bancaria, como las contraseñas, códigos de seguridad, etc., para cometer fraude.

Estos troyanos, comparten las mismas propiedades:

- Eluden los antivirus
- Interceptan las pulsaciones de teclado
- Almacenan los archivos
- Ayudan a los cibercriminales a husmear en las cuentas de las víctimas y a transferir su dinero ilegalmente.
- Intentan instalarse en los dispositivos móviles, para que los creadores accedan a los códigos de seguridad, que envían las entidades bancarias a través de SMS.

3.3.4. El problema de los IDS (sistemas de detección de intrusos)

Los IDS a menudo no son satisfactorios con respecto a las siguientes métricas:

- La detección basada en anomalías
- La detección basada en el mal uso o uso incorrecto

La detección basada en anomalías puede generar muchos falsos positivos (se considera que existe fraude de manera errónea), ya que una desviación del comportamiento normal especificado, no es necesariamente un ataque. Además, si la definición de comportamiento normal, se actualiza en tiempo de ejecución, un intruso experto puede cambiar lentamente su comportamiento, para finalmente incluirlo en la definición. Esto podría entonces dar lugar a un falso negativo (se considera que no existe fraude de manera errónea).

La detección basada en el mal uso o uso incorrecto, puede generar muchas alarmas perdidas, ya que para la mayoría de los sistemas abiertos de práctica es muy difícil definir una base de datos exhaustiva del ataque [YON2003]

3.3.5. El problema de la autenticación del usuario (PAU) en aplicaciones web

En [OWA2007] puede verse el ranking de vulnerabilidades en la red, donde se encuentra la vulnerabilidad "Broken authentication and session management" (el manejo de sesión y la autenticación infringida) en el séptimo lugar. Este problema sucede cuando las credenciales de cuentas y los tokens de sesión no suelen protegerse adecuadamente. Por ello, los atacantes comprometen contraseñas, claves o tokens de autenticación para asumir la identidad de otros usuarios. Esto perjudica al usuario, de manera que puede ocasionar un FTC sobre su cuenta bancaria en línea (al utilizar online banking).

El problema del manejo de sesión y de la autenticación infringida estaban en 2007 en el 7mo lugar (A7), entre los grandes problemas de vulnerabilidades de la seguridad web mundial.

En [OWA2010] nuevamente puede observarse en el top 10, que el problema de manejo de sesión y autenticación infringida escaló al 3er lugar (A3) en sólo 3

años. Luego, continúo escalando en los siguientes 3 años, hasta ubicarse en el 2do lugar [OWA2013]

Por la velocidad de crecimiento en el top 10 de [OWA2007] [OWA2010] [OWA2013], se puede afirmar que PAU es actualmente uno de los problemas de inseguridad web con mayor impacto en menor tiempo, respecto de otros problemas de inseguridad web actuales.

(Ver más en Anexo B)

Capítulo 4

Modelos computacionales para la DFTC

En este capítulo se describen los modelos computacionales, tales como, modelo de bloques de FDS, modelo de Hibridación BLAST-SSAHA, modelo oculto de Markov (HMM) y modelo de bloques del sistema evolutivo-difuso, que marcaron una tendencia en la resolución del DFTC.

4.1. Modelo de bloques de FDS

El sistema FDS consiste de cuatro componentes, llamados, rule-based filter, Dempster–Shafer adder, transaction history database y Bayesian learner.

En el componente rule-based, se determina el nivel de sospecha de cada transacción entrante, basada en el grado de su desviación de buenos patrones. Esto genera evidencias, que combina la teoría de Dempster–Shafer, para computar un concepto inicial. Luego, los valores del concepto inicial, son combinados para obtener un concepto general, aplicando la teoría de Dempster–Shafer. La transacción es clasificada como sospechosa o no sospechosa, dependiendo de su concepto inicial. Una vez que la transacción es considerada como sospechosa, el concepto es reforzado o debilitado, en función de su similitud con el historial de transacciones genuino o fraudulento, usando aprendizaje Bayesiano.

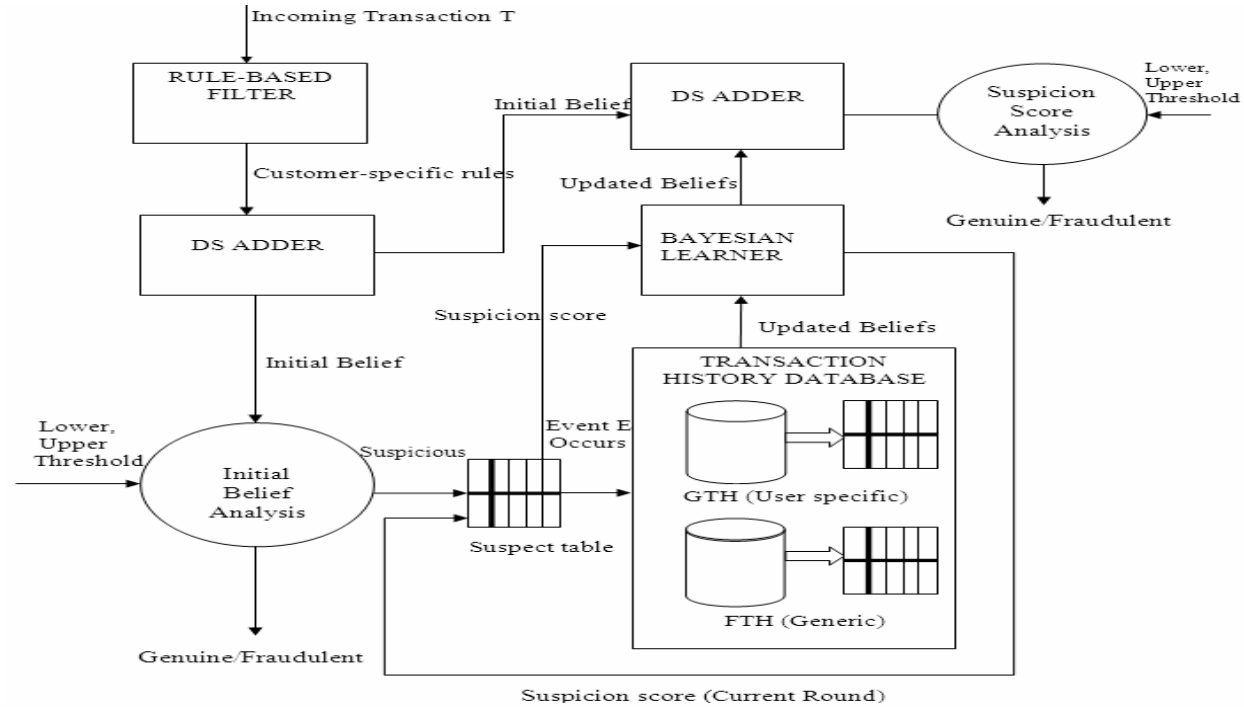

Figura 1. Modelo de bloques de FDS [EDW2011]

4.2. Modelo de hibridación BLAST-SSAHA

La hibridación de BLAST y el algoritmo de SSAHA es referido como el algoritmo BLAH-FDS. El alineamiento de secuencia resulta una técnica eficiente para el análisis del comportamiento de compra de los clientes.

BLAST y SSAHA son algoritmos de alineamiento de secuencia eficientes usados para la DFTC.

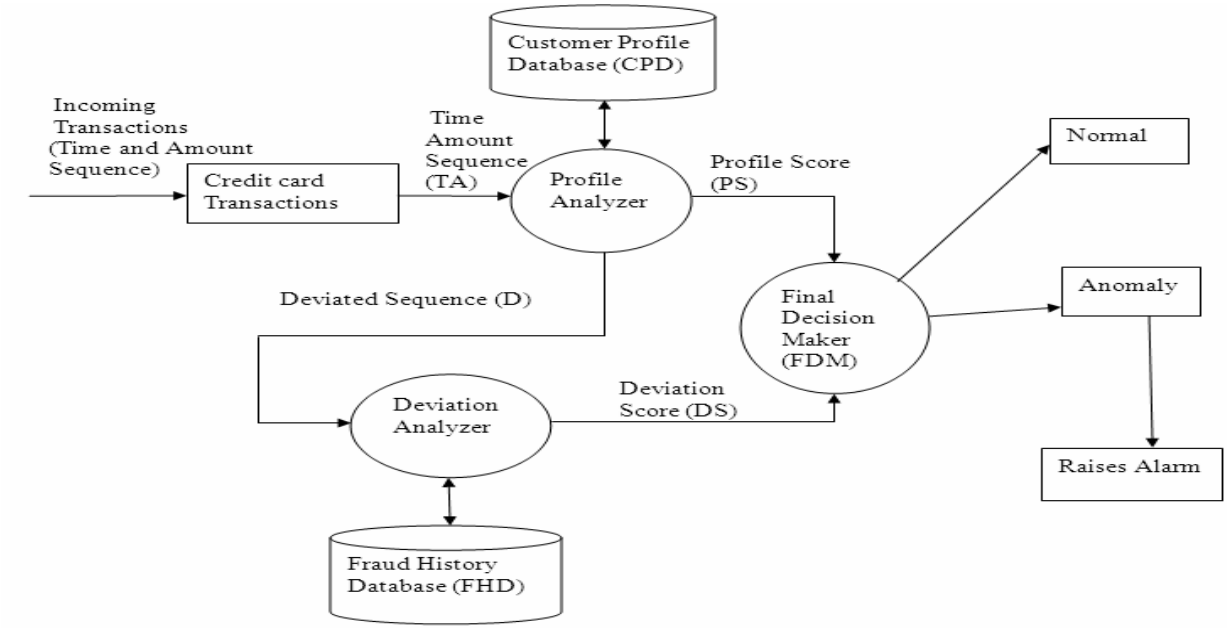

Figura 2. Arquitectura de BLAST y el Sistema de Detección de Fraude SSAHA [EDW2011]

BLAH-FDS es un algoritmo de alineamiento de secuencia de dos etapas en el cual un analizador de perfil (profile analyzer o PA) determina la similitud de secuencias de transacciones de entrada sobre una TC dada, con las secuencias de compras pasadas del portador genuino de la misma. Las transacciones inusuales trazadas por el analizador de perfiles (PA) son pasadas a un analizador de desviaciones (deviation analyzer o DA) para un posible alineamiento con el comportamiento fraudulento pasado. La decisión final sobre la naturaleza de una transacción, es tomada sobre la base de las observaciones por estos dos analizadores.

Cuando una transacción se lleva a cabo, la secuencia de entrada, se combina en dos secuencias de tiempo-cantidad, secuencia que se define como TA. La TA se alinea con las secuencias relacionadas con la TC en la base de datos del perfil del cliente (Customer Profile Database o CPD). Este proceso de alineamiento, es realizado usando BLAST.

El algoritmo SSAHA se usa para mejorar la velocidad del proceso de alineamiento. Si la secuencia TA contiene una transacción genuina, luego podría alinear bien con las secuencias en la base de datos del cliente (Customer Profile Database o CPD). Si existen algunas transacciones fraudulentas, pueden ocurrir desajustes en el proceso de alineamiento. Este desajuste, produce una secuencia desviada D, que está alineada con la base de

datos del historial de fraude (Fraud History Database o FHD). Una gran similitud entre una secuencia desviada D y FHD confirma la presencia de transacciones fraudulentas.

El analizador de perfil (PA), evalúa una puntuación de perfil (PS) de acuerdo a la similitud entre la secuencia TA y la base de datos del cliente (CPD).

El analizador de desviaciones (DA), evalúa una puntuación de la desviación (DS) de acuerdo con la similitud entre la desviación D y la base de datos del historial del cliente (FHD). La FDM (Final Decision Maker), finalmente, activa una alarma si la puntuación total (PS - DS) está por debajo del umbral de alarma.

4.3. Modelo oculto de Markov (HMM)

Un modelo oculto de Markov es un proceso estocástico integrado doble, que se utiliza para modelar procesos estocásticos mucho más complicados, en comparación con un modelo de Markov tradicional. Si una transacción de TC entrante, no es aceptada por el entrenado modelo oculto de Markov, con una probabilidad suficientemente alta, se considera como una transacción fraudulenta [**SRI2008**]

HMM, inicialmente, se entrena con el comportamiento normal del titular de una tarjeta. Cada transacción de entrada se somete al FDS para su verificación. FDS recibe los detalles de la tarjeta y el valor de compra, para verificar si la transacción es real. Si el FDS confirma que la transacción es maliciosa, activa una alarma, y la emisión del banco, rechaza la transacción. El titular en cuestión, puede entonces ser contactado y alertado, sobre la posibilidad de que la misma se vea comprometida.

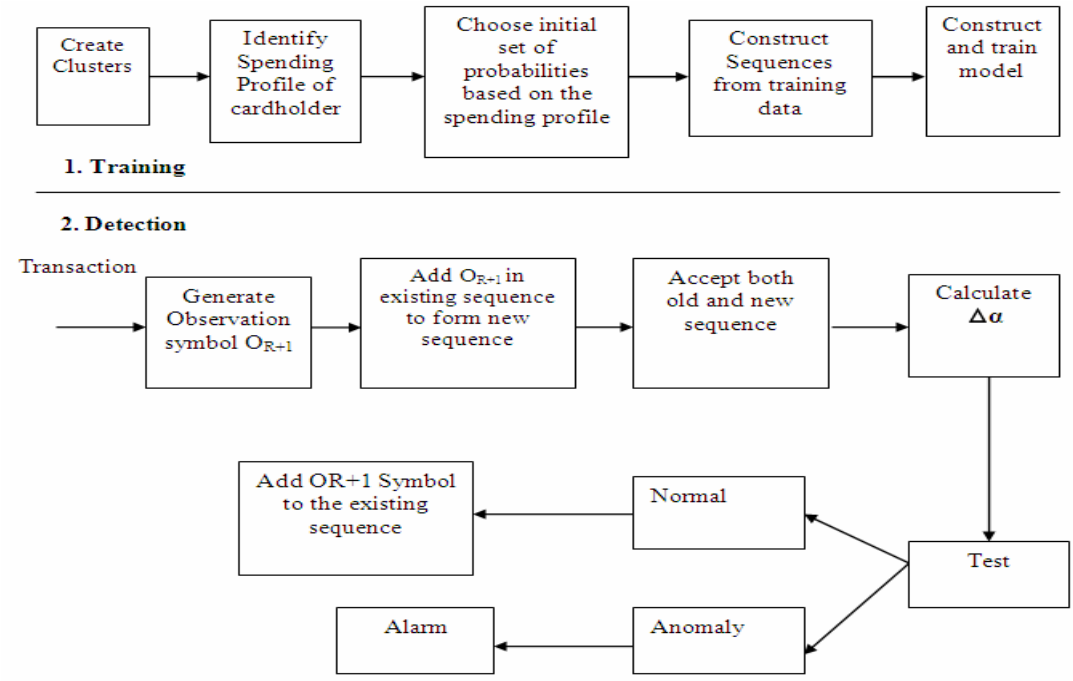

Figura 3. Modelo oculto de Markov- Flujo del proceso de FDS [SRI2008]

4.4. Modelo de bloques del sistema evolutivo-difuso

El Sistema de Detección Darwiniana Difusa [**BEN2000**] utiliza programación genética para desarrollar reglas de lógica difusa capaces de clasificar transacciones de TC en clases de "sospechosa" y "no sospechosa". Describe el uso de un sistema evolutivo-difuso capaz de clasificar transacciones de TC sospechosas y no sospechosas.

El sistema se compone de un algoritmo de búsqueda de Programación Genética (GP) y un sistema experto difuso.

Los datos se proporcionan al sistema FDS. Este sistema primero divide los datos en tres grupos, a saber, baja, media y alta. Los genotipos y fenotipos del sistema GP consisten en reglas que relacionan la secuencia entrante con la secuencia pasada.

La Programación Genética se utiliza para desarrollar una serie de reglas difusas de longitud variable que caracterizan las diferencias entre las clases de los datos contenidos en una base de datos. El sistema se desarrolló con el objetivo específico de detección de fraude, que consiste en la difícil tarea de

clasificar los datos en las categorías: "seguro" y "sospechoso". Cuando el pago del cliente no está vencido o el número de pagos atrasados es menor de tres meses, la operación se considera como "no sospechosa", de lo contrario, se considera como "sospechosa".

La Lógica Difusa Darwiniana detecta datos sospechosos y no sospechosos, y detecta fácilmente FTC robadas.

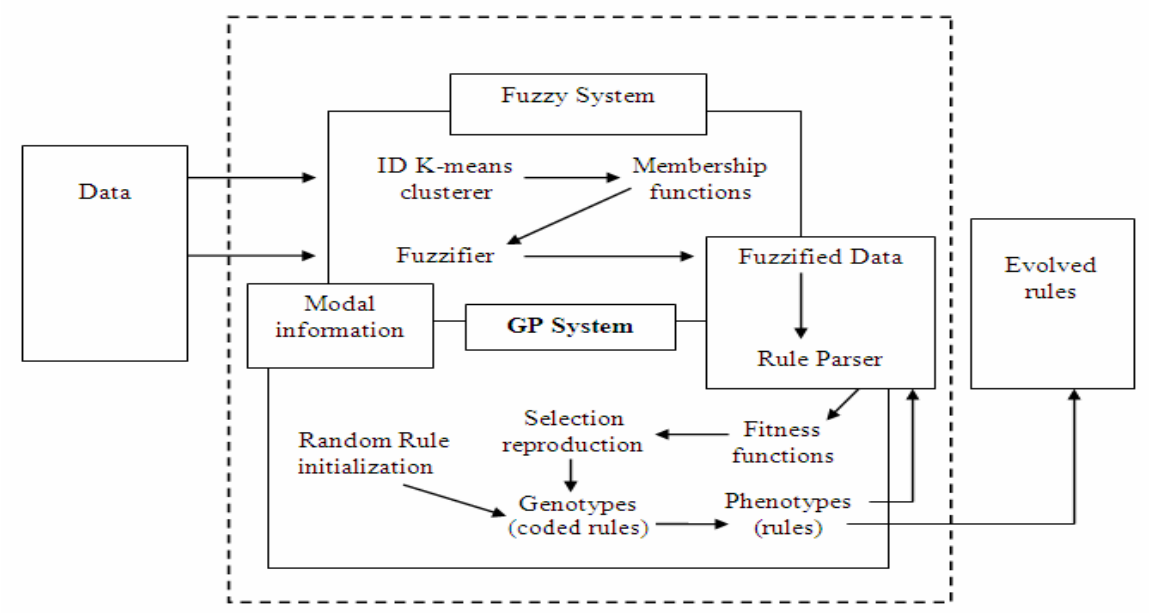

Figura 4. Diagrama de bloques del sistema evolutivo-difuso [BEN2000]

4.5. Análisis de ventajas y desventajas de modelos computacionales

El modelo de bloques de FDS, cuenta con una alta precisión y alta velocidad de procesamiento. Mejora la tasa de detección, reduce las falsas alarmas, y también es aplicable en E-Commerce. Por contraparte, es muy caro y su velocidad de procesamiento es baja.

A partir del 2011, este modelo cayó en desuso y no se presentaron mejoras.

En el modelo de hibridación BLAST-SSAHA, el rendimiento de BLAH-FDS es bueno y el resultado es de alta precisión. Al mismo tiempo, la velocidad de procesamiento es suficientemente rápida para permitir detección en línea de FTC.

Este modelo es capaz de detectar fraudes en telecomunicaciones y fraudes bancarios, pero no es capaz de detectar la clonación de TC.

A partir del 2011, este modelo cayó en desuso y no se presentaron mejoras.

El modelo oculto de Markov (HMM), nunca comprueba al usuario original, ya que mantiene un registro. Este registro que se mantiene, también es una prueba para el banco, por la transacción realizada. Por contraparte, produce muchas falsas alarmas y muchos falsos positivos.

A partir del 2008, este modelo cayó en desuso y no se presentaron mejoras.

En el modelo de bloques del sistema evolutivo-difuso, el sistema completo, es capaz de alcanzar una buena precisión, y niveles adecuados de inteligibilidad para datos reales. Cuenta con una precisión muy alta y produce pocas falsas alarmas, pero no es aplicable en las transacciones en línea y es muy caro. Por último, la velocidad de procesamiento del sistema es baja.

Este modelo se presentó por primera vez en el 2000, tuvo un único aporte en el 2008, luego, cayó en desuso, y no se presentaron mejoras, ni trabajos que lo citaran como referencia.

(Ver más en Anexo C)

Capítulo 5

Metodologías para la identificación del usuario fraudulento

En este capítulo se explica el uso de datos biométricos y el uso de métodos de ubicación para la identificación del usuario.

Existen diversas metodologías que se acercan gradualmente a una solución potencial del FTC. No se presentaron a la actualidad, modelos de detección viables, comprobables, que utilicen estas metodologías, pero apuntan a la detección del usuario fraudulento, que, como se ha visto en el capítulo 2, es la solución al problema central del FTC. Además, aporta una línea de investigación novedosa, netamente diferente, que amerita evaluar con detenimiento los avances presentados.

5.1. Uso de datos biométricos para la identificación del usuario

Según el centro de difusión de tecnologías de la Universidad Politécnica de Madrid, en la actualidad, existen sistemas que permiten la identificación de usuarios por las características únicas de una persona: algo que el usuario es. Estas características representan un patrón propio, que no puede coincidir con el patrón de ningún otro individuo, y que además, se considera difícil de reproducir. La biometría, tiene como objetivo, el estudio de las técnicas de reconocimiento de usuarios, utilizando las características corporales, que lo distinguen de otros usuarios.

La autenticación basada en características físicas, se utiliza diariamente de forma inconsciente, con el reconocimiento de personas, por los rasgos de su rostro o por su tono de voz. Cuando se utiliza un sistema biométrico, un dispositivo, reconoce estos rasgos. Este se basa en las características del sujeto a identificar, y luego, permite o deniega el acceso (a un determinado recurso o lugar físico) en respuesta a su evaluación previa.

Algunas características corporales que captan los sistemas de identificación biométrica son:

- Huellas dactilares.

- Patrón de las venas de la retina.
- Patrón del iris.
- Venas del dorso de la mano.
- Geometría de la mano.
- Rostro.
- Gestos.
- Patrón de voz.
- Firma manuscrita.

Existen dispositivos que disponen de sistemas de reconocimiento de huella dactilar incorporados. Estos permiten sustituir los nombres de usuario y contraseñas, y pueden, bloquear el disco duro para la protección de los datos.

En los sistemas de verificación de identidad, el individuo declara su identidad, y el objetivo, es averiguar si es quien dice ser. Para ello se comparan las características biométricas extraídas del individuo, con las almacenadas para ese individuo en la base de datos, y luego, se decide si concuerdan suficientemente. En caso contrario, se deniega el acceso.

Según el centro de difusión de tecnologías de la Universidad Politécnica de Madrid, se puede dividir a la autenticación en 3 características fundamentales:

1. Algo que el usuario sabe: una contraseña, un nombre de usuario, etc.
2. Algo que el usuario posee: una llave, una tarjeta, etc.
3. Algo que el usuario es: una característica corporal del mismo.

La biometría, utiliza la tercera característica. En las características 1 y 2, se puede observar, el riesgo de un eventual olvido del usuario (se puede olvidar una tarjeta, una contraseña o nombre de usuario). Pero al utilizar la tercer característica, no se corre el riesgo del olvido, pues, éste siempre lo lleva consigo (no se puede olvidar una huella dactilar, o el timbre de voz, por ejemplo). Y también, disminuyen considerablemente las posibilidades de falsificación.

Los sistemas de identificación biométrica se realizan en 4 pasos:

- Paso 1: captura
- Paso 2: extracción de datos
- Paso 3: clasificación

- Paso 4: decisión

En el paso 1, se obtienen aquellas características que se desean analizar. Mientras que, en el paso 2, se analizan o procesan las imágenes, sonidos u otros datos recibidos, para extraer ciertas características de individuos. Luego, en el paso 3, se alinean aquellas características extraídas junto con las almacenadas en una base de datos o archivo, y luego se comparan para obtener diferencias. Finalmente, en el paso 4, se reciben las comparaciones realizadas para tomar la decisión de autorizar el acceso, o denegarlo, en caso de que las características no coincidan.

Algunos ejemplos son:

- Reconocimiento facial: Una de las novedades de Android 4.0 Ice Cream Sandwich, es la capacidad de desbloquear el teléfono con el rostro. Es una aplicación de la tecnología de reconocimiento facial práctico y destacable, pero también trae consigo algunos problemas de seguridad, ya que, con una simple imagen delante del terminal, se puede engañar al sistema y dar paso a una gran cantidad de información que tenemos depositada en el Smartphone.

Figura 5. El desbloqueo facial de Android 4.0 Ice Cream Sandwich

- Reconocimiento facial y múltiples usuarios: Se puede observar en dispositivos iOS de Apple, el uso del reconocimiento facial, pero, con la característica de ser multiusuario. Esto permite que varias personas utilicen el mismo aparato, sin que la privacidad de las otras personas se

vea afectada, por eso, los mensajes, emails, fotos, etc. de un usuario, no están disponibles para otro usuario.

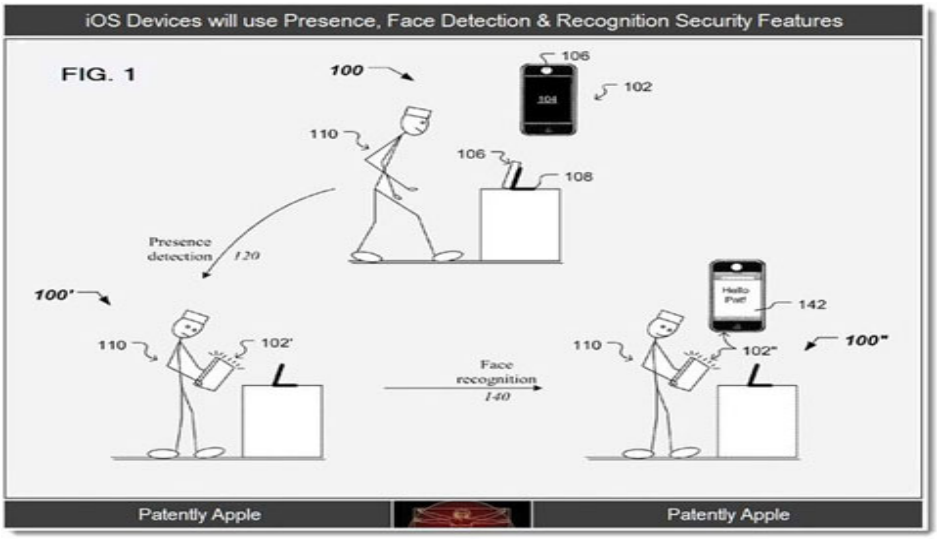

Figura 6. El sistema de reconocimiento facial y múltiples usuarios en dispositivos iOS de Apple

- Reconocimiento de objetos en tres dimensiones: Apple acaba de conseguir registrar una nueva patente, entre las que encontramos el reconocimiento de objetos en tres dimensiones. Como se dijo anteriormente, Android 4.0 Ice Cream Sandwich, puede reconocer caras y desbloquear un terminal automáticamente, al ubicar el rostro delante del terminal. Ésta patente de reconocimiento de objetos en tres dimensiones, tiene como objetivo esta misma funcionalidad, lo que demuestra que en Apple, están trabajando para conseguir un sistema similar, pero que funcione sólo con caras de verdad, en lugar de fotografías planas. Esto podría evitar situaciones de fraude a través del reconocimiento de un rostro real.

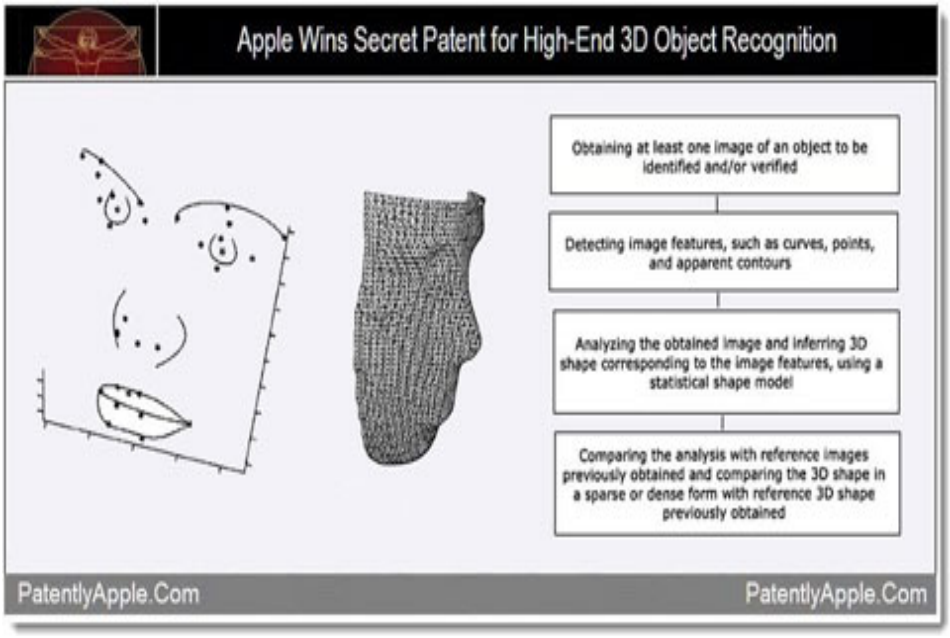

Figura 7. Sistema de reconocimiento de objetos en 3D de Apple

- Lectura de patrones de las venas de la palma de la mano: Una técnica biométrica muy novedosa es la de Fujitsu Services, que presentó un kit personal de seguridad biométrica "PalmSecure Kit". Este integra un sensor para la lectura de patrones de las venas de la palma de la mano, y un software de seguridad para gestionar el acceso al PC y a las aplicaciones, lo que supone un avance muy importante en la seguridad del acceso al mismo. Las venas de la mano, al estar a dos o tres milímetros por debajo de la epidermis, y ser por tanto internas a nuestro cuerpo, no se pueden copiar. Además, el patrón de las venas es único en cada individuo, incluso en el caso de gemelos idénticos, así mismo, son diferentes las venas de la mano derecha, y de la izquierda. Tampoco cambia con el crecimiento, sólo se amplía, pero manteniendo el mismo patrón. Esta tecnología, se está adoptando en entidades financieras, bancarias, y ATM en todo el mundo.

Figura 8. Dispositivo lector de las venas de la palma de la mano de Fujitsu [FUJ2014]

5.2. Uso de métodos de ubicación para la identificación del usuario

Una forma novedosa de identificar si el usuario es quien dice ser, es conocer la ubicación del dispositivo y verificar lo que suele llamarse como "las ubicaciones de familiaridad". Esta idea, refiere a que el usuario, suele moverse con el terminal en ubicaciones conocidas y reincidentes, por lo que, se conserva en el dispositivo, un historial de ubicaciones. Luego, se consiguen las coordenadas actuales, bajo diferentes métodos. Finalmente, se decide si el terminal, se encuentra dentro de ese rango de familiaridad con el que el usuario suele moverse, para habilitar o deshabilitar aplicaciones en uso, o el terminal mismo.

Para localizar al usuario, se debe localizar el dispositivo, de modo que, las metodologías utilizadas para la detección del dispositivo, detectan un usuario. Si un usuario tiene por finalidad, obtener datos de una TC, o ingresar a una aplicación bancaria ubicada dentro del terminal, esta metodología mencionada, colabora entonces, con la detección de un usuario fraudulento en FTC.

Los métodos para detectar la ubicación del dispositivo se realizan a través de:

- GPS
- Wi-Fi
- Torres de telefonía GSM

En el exterior, la detección de la ubicación está suficientemente provista por GPS, sin embargo GPS no es adecuado para el desafío de los ambientes interiores sin línea de visibilidad directa. En estos entornos de menor escala deben ser empleadas, técnicas de estimación de localización. Debido a su ubicuidad, las señales WiFi son un indicador comúnmente empleado para la ubicación; el conocimiento de la identidad y la intensidad de estas señales a lo largo de un entorno, pueden permitir la estimación de la ubicación del dispositivo de recepción [KEL2009]

En [CAB2003] puede observarse un caso de estudio, que utiliza torres en tierra, para conseguir la ubicación del dispositivo móvil, y en base a esa ubicación, ofrecer servicios. Este trabajo está centrado en el problema del acceso a la información, y servicios, relacionados con los recursos turísticos de una ciudad, con el objetivo de organizar adecuadamente una visita "on-the-fly" o sobre la marcha. Supone también, que la ciudad cuenta con una infraestructura de comunicación distribuida adecuada. En particular, que la ciudad contiene "torres de información", distribuidas por la ciudad, y que suministran información, y servicios relacionados con el lugar específico a la región local de la ciudad, en el que se encuentra el usuario (ver figura 9). A modo de ejemplo, se debería encontrar tal infraestructura en Roma (Italia), luego, una torre de información por parte de la Ciudad del Vaticano proporcionaría información acerca de visitar los museos del Vaticano y la Cappella Sistina; una torre de información de Trastevere, proporcionaría información sobre la Iglesia de Santa María y servicios para reservar una mesa en un restaurante típico de Trastevere. Estas torres de información, materializan claramente la abstracción de contextos de servicio local. El acceso a la información y a los servicios de las torres de la información, se realiza a través de diferentes medios. En primer lugar, las torres de información están habilitadas para ofrecer información y servicios a través de conexiones inalámbricas (caso 1, en figura 9). Así, los turistas que están actualmente visitando la ciudad, moviéndose en sus calles, pueden explotar un asistente personal digital, que se ejecuta en un ordenador de mano (palm computer), o teléfono inteligente, para acceder a las torres locales de información, y allí descubrir lo que hay alrededor. Por ejemplo, una persona que camina en Trastevere, se puede conectar a la torre de la información local, para descubrir si hay un restaurante chino en el barrio, y posiblemente, para reservar una mesa. Por otra parte, se supone que las torres de información están conectadas a Internet, de modo que puedan acoger la ejecución de agentes de software móvil. Los usuarios pueden así, enviar a sus agentes móviles, asistentes personales, para recorrer a través de las torres de

información, y allí, acceder a información y servicios. Esta instalación, se podría utilizar para enviar agentes móviles, para organizar una visita fuera de la línea de la ciudad, antes de la actual (caso 2, en figura 9). También, una vez que el turista está en un lugar, esta instalación podría ser utilizada para desplegar un agente móvil en la red, para acceder a torres de información remota, sin tener el turista que estar caminando por ese lugar (caso 3, figura 9).

El escenario anterior puede ser considerado como representante de varias otras áreas de aplicación, como los hogares inteligentes, sistemas de información de tráfico, o los museos inteligentes. De hecho, todas estas áreas de aplicación, se caracterizan por la presencia de una infraestructura fija, sobre la base de una multiplicidad de nodos, cada uno asociado a una ubicación específica, que proporciona información sensible a la ubicación y servicios.

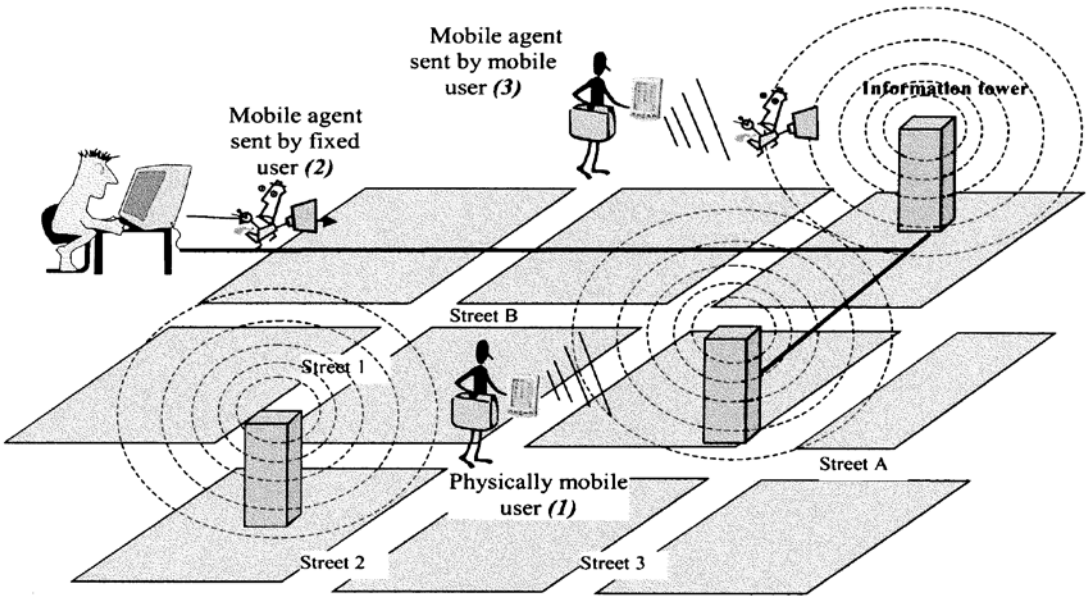

Figura 9. Ubicación por torres de telefonía

5.3. Análisis de ventajas y desventajas de las metodologías propuestas

Respecto al uso de datos biométricos para la identificación del usuario, y según el centro de difusión de tecnologías de la Universidad Politécnica de Madrid, se observan las siguientes ventajas y desventajas.

Uno desventaja con las aplicaciones biométricas, es que es necesario disponer de una base de datos centralizada, donde residan los patrones correspondientes a los datos biométricos de los usuarios. Estos datos, son personales y privados, por lo que deben estar regulados y protegidos a nivel nacional e internacional. Además deben ser almacenados en un servidor seguro, puesto que el acceso no permitido a esa base de datos, invalidaría completamente la operativa del sistema.

En los móviles, la autenticación biométrica se realiza utilizando el hardware instalado en el terminal móvil, y los datos se almacenan en dicho dispositivo, sin necesidad de recurrir a una base de datos externa centralizada. Los datos biométricos se capturan, procesan y almacenan de forma segura en el propio terminal.

Son ya muchas las compañías y organismos, que empezaron a experimentar las posibilidades que tiene la utilización de sistemas de seguridad basados en la biometría, y en uso de diferentes sensores. Esto ofrece una nueva manera de identificarse, basándose en lo que es la persona, utilizando algo que forma parte de su cuerpo, de su propia identidad.

Además de seguridad, ofrece comodidad, ya que permiten que el usuario, prescinda de llevar tarjetas consigo, o de tener que acordarse de las contraseñas o claves que le dan acceso a los sistemas o emplazamientos. Poner el dedo en un sistema de captación de huellas digitales, mirar a un dispositivo de reconocimiento de iris o hablar, así como presentar su ubicación a través del dispositivo, es lo único necesario para poder autenticarse a través de los diferentes dispositivos de captación de patrones biométricos.

Por otro lado, la característica biométrica que se puede utilizar, depende de varios factores: los sistemas disponibles en el dispositivo móvil, la capacidad de almacenamiento en memoria y la capacidad de procesamiento.

En los ejemplos mencionados, se observa la tecnología de reconocimiento facial, como novedosa y destacable, pero a la actualidad, el mismo sistema puede ser engañado con una simple imagen digital o fotografía del usuario, y por ello, dar acceso, generando lo que se conoce como falsos negativos sobre la identidad. Una vez engañado el sistema, el usuario tiene acceso a una gran cantidad de información del Smartphone, y esto la hace propicia para el FTC.

Se está trabajando para conseguir un sistema que funcione sólo con rostros reales, en lugar de fotografías planas. Esto evitaría situaciones de fraude por reconocimiento de rostro real.

Desde otro punto de vista, se observa que todavía se basa en la seguridad del dispositivo mismo, y no en la seguridad de cada aplicación móvil. Por esto, una vez ingresado al dispositivo, ya no existen restricciones de acceso a ninguna aplicación. Se espera que en trabajos futuros se refuerce la seguridad en las aplicaciones, y que las soluciones se combinen con este tipo de metodologías. No sólo, con el ingreso de usuario y contraseña, sino apuntando a la identificación del mismo, por lo que el usuario es.

La mayoría de las patentes sobre reconocimiento de rostro y objetos 3D en dispositivos móviles, se obtuvieron en el 2014.

En cuanto a la tecnología de lectura de las venas de la palma de la mano, Fujitsu asegura que PalmSecure ® proporciona una exactitud en autenticación líder en la industria, con tasas de falsos resultados extremadamente baja, y el dispositivo de lectura, sin contacto, higiénico, y no intrusivo, proporciona facilidad de uso, prácticamente, sin restricción fisiológica para todos los usuarios.

Así, seguridad, comodidad y rapidez, son tres rasgos propios de la nueva era de seguridad móvil, que viene dotada de tecnologías para la identificación del usuario a través de datos biométricos.

Respecto al uso de métodos de ubicación para la identificación del usuario, se observa las siguientes ventajas y desventajas.

El mayor incentivo actual para la investigación sobre métodos de ubicación de dispositivos, es ofrecer servicios relacionados con el contexto, tal que existen aplicaciones “context aware” que se benefician de estas investigaciones. Pero, existe también, una preocupación por ofrecer resguardo a clientes, de servicios que se realizan a través de aplicaciones que pueden sufrir fraude a la identidad, para obtener beneficios económicos. De allí que, los proyectos que se presenten con la idea de mejorar la ubicación del usuario, naturalmente, mejoran las posibilidades de identificar a este usuario, al chequear si este se encuentra en una ubicación de familiaridad para el dispositivo.

Esta es una alternativa novedosa y precisa, pero no se observan trabajos actuales, que se basen sólo en la identificación del usuario. Se espera que se presenten trabajos, ligados exclusivamente a la seguridad del dispositivo, y a las aplicaciones que se apoyan sobre el mismo.

Finalmente, se debe destacar, que los trabajos presentados con orientación turística, enriquecieron el material disponible en cuanto a información del contexto del usuario. Esta información, sirve de soporte para la identificación del usuario, por su ubicación y por las características del medio donde realiza operaciones.

En trabajos presentados, como en [FIS2012], se observa una interesante forma de dividir la seguridad, en niveles de seguridad, de acuerdo a la ubicación del usuario.

(Ver más en anexo D)

Capítulo 6

Propuesta de modelos futuros

En este capítulo se describen los pasos necesarios para lograr modelos futuros sustentables que prevengan el FTC. Se describe la definición de un usuario fraudulento, la apertura de caminos viables, la definición de la forma de trabajo y sus restricciones, luego se define un modelo computacional aplicable de ejemplo, para finalmente tratar, la implementación del modelo y las pruebas estadísticas sobre diversos usuarios. Por último, se muestra un modelo computacional genérico estimativo para la DFTC.

6.1. Características a mejorar en nuevos modelos

En base a los temas tratados en capítulos anteriores, se puede afirmar que las siguientes características, se deben tener en cuenta para generar modelos futuros exitosos.

- **Definir en detalle, qué se considera un usuario real, y qué es un usuario fraudulento.**

  Queda demostrado que se desconocen las características de un usuario fraudulento; y al incrementarse día a día la adhesión de nuevos y variados usuarios, este problema, es un problema creciente.
  Algunas aplicaciones ya reconocen quién es el usuario titular, pero no reconocen aun quién no lo es. Esto se reveló en capítulos 3, 4 y 5.

  En el capítulo 3, queda demostrado que la autenticación ingresando nombre de usuario y contraseña, no es suficiente para identificar al usuario titular. Más aún, si un usuario fraudulento logra conocer estos datos, puede delinquir sin ninguna restricción. Esto se debe a que, no se identifica al usuario por lo que es, sino por los datos que este conoce. Según esta teoría, todo usuario que conoce los datos de otro usuario, es considerado un usuario titular. Cuanto más fácil es descubrir esos datos, más usuarios titulares se tienen para un sistema de detección. Entonces, se tienen más usuarios fraudulentos.

  En el capítulo 4, cuando se proponen varios modelos de FTC que guardan en un historial el comportamiento del usuario y sus

transacciones, en la mayoría de los trabajos propuestos, se indica el problema de los falsos negativos. Estos sistemas detectaban al usuario titular, pero también aceptaban a un usuario fraudulento como un usuario titular, cuando no lo era. Tampoco se encontró trabajos posteriores que demostraran haber encontrado la solución a este problema. En estos trabajos, puntualmente, se demostró que el comportamiento del usuario, generalmente no representa quién es el usuario real, o al menos, no es suficiente para develar su identidad. Dado el caso de utilizar este concepto, este debe combinarse con otros.

En el capítulo 5, se observa un caso en el que una foto del usuario titular, permite el acceso al dispositivo móvil. Esto sucede porque, la aplicación reconoce un rostro de manera planar y no en 3D, y además, porque evalúa quien es el usuario titular, pero no distingue quién no lo es.

En los modelos futuros, se debe definir, qué es un usuario fraudulento, y modelarlo en detalle, en cada etapa del sistema en el que se opera. También deben considerarse, otros métodos de identificación, no sólo basados en el ingreso de usuario y contraseña. Puntualmente, no sólo basados en la identificación del usuario, por la información que conoce.

- **Incluir la posibilidad de fallo en la autenticación inicial.**

  Los usuarios fraudulentos acceden por la etapa 1(infringiendo la seguridad online), pero principalmente, logran acceder a la etapa 3 (el usuario ya se encuentra operando en el sistema), y cada vez con menor dificultad. O sea, que traspasan la autenticación del usuario clásica, elaborada en base a perfiles de usuarios, donde se limita el uso de diversas aplicaciones y operaciones de estas, de acuerdo a cada perfil. Por esto, es que debe considerarse la posibilidad de implementar una autenticación por operación y/o por aplicación, teniendo en cuenta el posible acceso a la etapa 3. Entonces, si se mantienen los perfiles, debe considerarse la posibilidad de que un usuario con perfil de operador, puede conocer los datos para ingresar al perfil del administrador, y una vez identificado en el sistema, este puede realizar operaciones, como un alta, una baja, o una modificación de datos, o administrar la configuración. O sea, que tiene libre acceso a las operaciones del perfil administrador. En estos casos, puede plantearse el hecho de pedir autenticación o chequear la autenticación, asociada a la operación

misma. Se puede hacer un seguimiento o una doble autenticación del usuario registrado, siempre que no genere demoras innecesarias.

Esta idea, refiere al hecho, de no asumir que el usuario ingresado, realmente es quien dice ser. Por lo que, se debe continuar controlando a ese usuario, no sólo con un tiempo límite de sesión, sino, buscando otra forma de identificación, durante la sesión.

- **Generar modelos adaptables a dispositivos móviles.**

  Cada vez más operaciones se realizan a través de aplicaciones móviles, y las operaciones que resulten en fraude, no serán la excepción. Por esto, los modelos futuros propuestos para la DFTC, deberán ser aplicables en varios tipos de dispositivos móviles. Aquí puede aplicarse la identificación por lugares de familiaridad evaluados en el capítulo 5, donde la ubicación del dispositivo colabora con la ubicación del usuario fraudulento.

- **Poner énfasis en modelos más que en metodologías.**

  Los modelos y metodologías propuestos a la fecha, han evolucionado bastante, pero todavía no resuelven el problema del FTC. Se deben proponer al menos nuevos modelos, que planteen el uso de las metodologías vistas en el capítulo 5 u otras vigentes y futuras.

  Si bien las metodologías que resuelven el problema de la falsa identidad, se han incrementado y son variadas, no se exponen trabajos con modelos para estandarizar soluciones (a excepción de los vistos en el capítulo 4). Esto implica, dos problemas graves.

  El primero, es asumir que un usuario es sólo algo que conoce o alguien que se comporta de manera similar a lo esperado. En particular, porque se estandarizó esta forma de identificación, y todos los sistemas actuales lo utilizan. Por esto, los modelos que se presenten, pueden comenzar a adoptar nuevas definiciones del concepto de usuario, y estandarizar algunos modelos computacionales. Esto puede generar una colaboración, entre diversos grupos de investigación, que utilizan diversas metodologías, pero para un mismo modelo computacional, al menos abstracto o genérico.

El segundo, es asumir que toda la carga de la identificación, esté dada por la metodología a utilizar. Principalmente, porque como paso previo a utilizarlas, se debe debatir con modelos computacionales, los conceptos de identificación del usuario que pueden coexistir. Porque puede darse el caso, por ejemplo, de que si se identifica a un usuario por el patrón de venas de la mano, ya no sea necesario, y resulte costoso, también identificarlo por el patrón de la retina de los ojos. Este tipo de problemas, pueden evitarse si se evalúan previamente, modelos computacionales estándar.

6.2. Pasos a seguir en próximos modelos de DFTC

Para poder realizar modelos futuros, se deberán aplicar los siguientes pasos.

6.2.1. Definición de un usuario fraudulento

En diferentes trabajos se ha observado que la definición de lo que es y lo que no es, un usuario fraudulento, no es una tarea trivial, en particular, porque el rango de usuarios es cada vez más variado. Personas de diferentes niveles de conocimiento, edad, entre otras características, trabajan, realizan transacciones bancarias y operan sobre el e-commerce, comprando boletos de vuelo, artículos variados y demás. Por ello, es una tarea compleja definir el usuario real, y el usuario fraudulento.

Para que un sistema reaccione ante el FTC, primero debe saber qué se entiende por usuario fraudulento, y para ello, deberán tenerse en cuenta los siguientes conceptos que podrán usarse en la definición, tal que estos conceptos están presentes en cada usuario.

Concepto 1: “El cuerpo no sabe mentir”:

- Por gestos habituales del usuario titular (posición de la comisura de los labios en tiempo, movimientos de izquierda a derecha, de arriba hacia abajo, risa continua, etc.)
- Por expresiones del rostro continuas del usuario detectado (fruncir cejas, apretar labios, etc.)

- Por estados de ánimo, en particular de sobresalto (Estado de las pupilas, sudoración de rostro y manos, levantamiento pronunciado de los hombros, etc.)
- Por aspecto (tipo, largo y color de pelo, rostro con y sin maquillaje, distinciones en el rostro como cicatrices, lunares, nariz pronunciada, surcos en la piel, colores pronunciados en surcos y líneas de expresión, etc.)
- Por datos biométricos (patrón de las venas de la mano, patrón de venas de la retina, etc.)

Concepto 2: "Siempre vuelvo a lugares conocidos":

- Por ubicaciones de familiaridad (aeropuerto, centro de la ciudad, países frecuentes, mi casa, etc.)

Concepto 3: "Regularmente me contacto con…":

- Por contactos más frecuentes (madre, amigos, compañeros de trabajo, etc.)

Concepto 4: "Regularmente realizo la misma operación"

- Por periodos recurrentes a la operación (1 vez al mes, 1 vez por semana, cada 2 semanas)
- Por rangos horarios comunes u horas pico.

6.2.2. Apertura de caminos viables

Si la idea del usuario fraudulento es, por ejemplo, la de "El cuerpo no sabe mentir", entonces deberá especificarse las formas de trabajo posibles como sigue:

- Uso de interfaces gestuales en tiempo real y comparación en tiempo real
- Grabación de gestos para su inmediata comparación.
- Captura de imágenes para su inmediata comparación.
- Grabación de lectura de frases para comparar la voz.
- Etc.

6.2.3. Definición de una forma de trabajo y sus restricciones triviales

Si se elige, por ejemplo, la comparación de imágenes captadas contra historial de imágenes, entonces:

- Se debe contar con al menos una imagen en el historial
- Las imágenes del historial por el momento se recomienda que se capten en un ambiente seguro, al solicitar la TC en el banco, por ejemplo, para evitar posteriores inyecciones a la BD fuera del lugar o fecha, entre otros.

Si el caso planteado es el ejemplo del 1er paso: "El cuerpo no sabe mentir", evaluando los "estados de ánimo", entonces:

Debe tenerse en cuenta factores externos como, el clima donde se encuentra el usuario. Si la temperatura es alta, las manos sudan naturalmente en el usuario, este usuario, no se encuentra entonces, bajo una situación de nervios o presión. Consecuentemente, la prueba debe ser descartada ante un hecho trivial. Si los factores ambientales o de contexto ponen en peligro la veracidad de las pruebas, entonces la definición del usuario fraudulento dada, no sirve para futuras pruebas. De allí, que debe ser descartada desde la definición misma.

Si la persona posee algún "tic", como levantar los hombros repentinamente, entonces tampoco podrá considerarse como un dato de alarma, dado que es un acto involuntario, por trivialidad se descarta como información.

Volviendo al ejemplo presentado, si se observa el brillo de la sudoración de rostro y manos, nos encontramos ante una búsqueda que puede dirigirse a encontrar brillo excesivo en la imagen capturada o cuadro de video (capturado en puntos cruciales a evaluar), y luego definir un resultado, puntualmente, si ese usuario es fraudulento o no.

En varios trabajos presentados se observa un avance importante respecto al reconocimiento de rostro, y la comparación de imágenes o cuadros de video en áreas localizadas. Por ejemplo, se compara la nariz y las distancias entre facciones, en vez de comparar la imagen completa. También se ha logrado una

optimización de los tiempos en trabajos de aplicaciones en tiempo real, tanto para comparación de rostros en imagen, así como audio y video.

6.2.4. Definición de un modelo computacional aplicable al paso 6.2.3

A. Definición del usuario fraudulento.

El usuario fraudulento tomado como ejemplo para desarrollar este punto, se basará en el estado de ánimo del usuario, chequeando dos simples características:

- Brillo excesivo en rostro y manos
- Estado de sobresalto captado en las pupilas

B. Datos del modelo.

- **Tipo de usuario a probar:** Fraudulento
- **Solicitud:** Transferir dinero de mi cuenta a la cuenta de mi amigo Pérez
- **Nivel de seguridad requerido:** Alto
- **Requerimiento de la operación:** Captar imagen cuerpo entero
- **Evaluación:** Brillo en rostro y manos, estado de sobresalto en pupilas
- **Datos del ambiente.**
  Clima: Fresco
  Hora: 20hs
  Tipo de aplicación: Móvil

C. Estado del chequeo interno

- Chequeando brillo en rostro y manos. Resultado: brillo excesivo, positivo
- Chequeando estado de las pupilas. Resultado: sobresalto, positivo

D. Mensaje al usuario:

- Positivo: Fraude
- Por favor contáctese con su banco. Cuenta cerrada. Apagado.

E. Acciones consecuentes:

- **Resultado positivo:** Bloqueo del equipo, email u otro mensaje al usuario real, como por ejemplo, 'Intento de fraude a su cuenta, comuníquese con su banco, cuenta bloqueada'

- **Resultado negativo:** Autorización de la transferencia de dinero. Se debe indicar nombre del contacto receptor de la transferencia. En este nivel, también puede adherirse otra capa de seguridad por contacto.

6.2.5. Implementación y pruebas estadísticas sobre diversos usuarios.

Lo más importante, es hacer foco en las características definidas como usuario fraudulento, la captura tomada como fuente de comparación, y si efectivamente, el resultado es positivo o negativo, cuáles serían las acciones a tomar. Una vez obtenido esto, se debe contar con la tecnología de prueba, y evaluar los tiempos de ejecución, falsos positivos, falsos negativos, etc., para ajustar la aplicación resultante. En algunas metodologías tratadas, todavía no se dispone de tecnología de soporte para realizar pruebas.

Se realiza el entrenamiento de la herramienta sólo si es necesario, esto sería el caso de elegir una línea basada en comportamiento del usuario mediante aprendizaje.

6.2.6. Modelo general de cómputo para la DFTC

En el siguiente modelo, puede observarse a un usuario fraudulento que sigue un modelo orientado a la interacción con el contexto a través de una aplicación móvil bancaria. Una vez que el usuario elije la operación, se describen los pasos que sigue el sistema para completar la operación. Puede verse la estructura general y el camino elegido en el ejemplo 6.2.4, este es:

1. "Usuario A" es un usuario fraudulento que ingresa en una aplicación bancaria desde su dispositivo. Este conoce el nombre de usuario y contraseña del titular.
2. Usuario A "Gana acceso" (en la autenticación por ingresar nombre de usuario y contraseña correctas). La autenticación se considera fallida.

3. El usuario A selecciona la operación “Transferir dinero a”. Tal que desea transferir dinero a su amigo Pérez.
4. La aplicación refuerza la seguridad al pedirle una segunda autenticación, tal que le pide que ingrese una foto actual tomada desde el dispositivo.
5. Implícitamente se llama a “Chequear brillo en manos y rostro, y estado de las pupilas”.
6. La aplicación detecta fraude, porque el usuario se encuentra sobresaltado y confuso. Por ello, se envía un mensaje al usuario, para que se comunique con su banco, cierra la cuenta y apaga el teléfono.

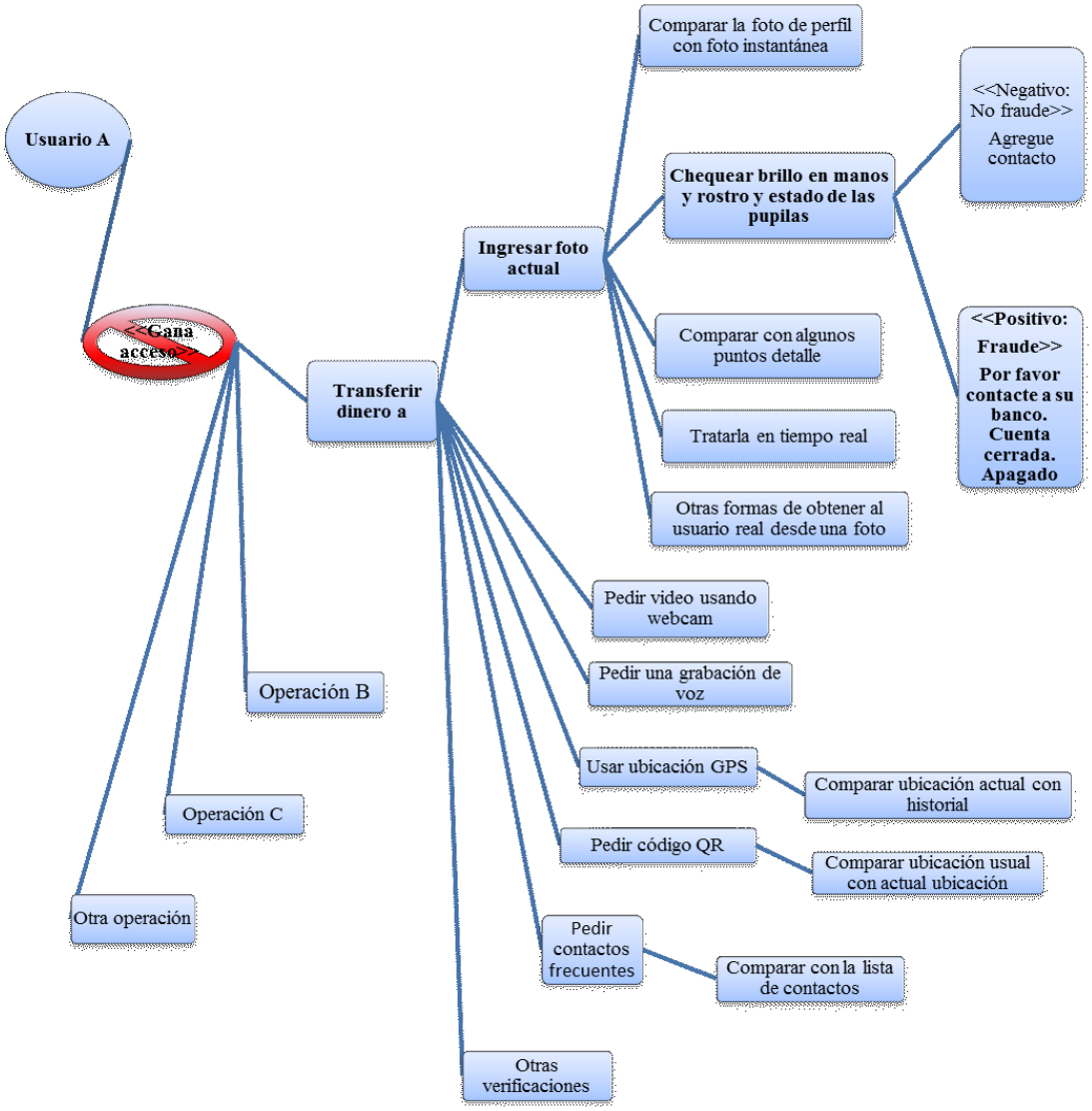

Figura 10. Modelo general de cómputo para la DFTC

Como se puede observar, el detectar exitosamente al usuario fraudulento o admitir su acceso, depende totalmente de su definición.
En este modelo también se puede observar, lo que sucede cuando se gana acceso (ver capítulo 2, etapa 3), y se puede partir de allí mismo, para generar otro tipo de modelos, ya no tan dependientes del ingreso del nombre de usuario y contraseña, ni de la autenticación inicial solamente.

Capítulo 7

Líneas de investigación encontradas en la DFTC

En este capítulo se describen las líneas de investigación de la DFTC encontradas. Para comenzar, se describe la línea de la seguridad en la red, y en el acceso a los datos, continúa con la línea de extracción de información del contexto, y finalmente, la línea basada en aprendizaje.

7.1. Línea de investigación relacionada con la seguridad en la red, y en el acceso a los datos

En el capítulo 3, se describió el esfuerzo de esta línea de investigación, por proteger tanto la información que viaja por la red, la red misma, así como el acceso a los datos del usuario. Se observó el hecho de que, a través de una red, mediante una aplicación web, email o sms, un usuario es engañado para acceder a enlaces falsos. También se notó que, se pueden generar descargas de software malicioso, en pos de exponer ciertos datos de la TC, así como datos personales, para cometer un fraude. El usuario también puede ser engañado, con el fin de obtener el acceso remoto a su computadora de escritorio o dispositivo móvil. Se percibe además, cómo se engaña a un usuario, para que realice acciones sin su conocimiento y sin su consentimiento.

Esta línea pretende detectar actividad fraudulenta en la red, analizando el tráfico en la misma, asegurando el uso de protocolos para la comunicación eficaces, de vanguardia, encriptando paquetes de información enviados, evitando la instalación de software malicioso, con actualizaciones del SO, y en algunas aplicaciones, generando una identificación segura a través de codificaciones de máquina o de usuario, en esa conexión.

Una ventaja de esta línea, es el hecho de poder recaudar cierta información (hardware o software) del atacante y de la fuente del ataque. Esta información, bajo otras líneas, a veces no es posible obtener. Como por ejemplo, IP del atacante.

Otra ventaja refiere, a la probabilidad de fallos. Se sabe que bajo otras líneas, el diagnóstico de fraude puede ser sólo un fallo en el diagnóstico del sistema,

en cambio, en esta línea, dado el caso de detección de un intruso en la red, puede asegurarse que realmente se trata de un ataque. Con lo cual, se puede decir que la probabilidad de fallos es baja o nula.

Se describió en el capítulo 2 y se demostró en el capítulo 3, el problema actual del acceso a la etapa 3. En esta etapa, el usuario gana acceso, por conocer el nombre de usuario y contraseña, requerido en la autenticación clásica. Es especialmente en esta temática, donde esta línea presenta varias deficiencias.

Otra desventaja es que, no se presentaron trabajos de esta línea que intenten identificar al usuario por lo que este es, sino sólo por la información que este conoce. Como se vio en el capítulo 5, la extracción de información de contexto, puede colaborar con esta tarea.

Aunque siempre se debe brindar un nivel de seguridad aceptable en las conexiones establecidas, como se denotó en los rankings de OWASP, existen formas de traspasar ese nivel, para acceder mínimo a la etapa 2, y según OWASP inclusive a la etapa 3. Esto sucede, cuando el usuario fraudulento, obtiene los datos suficientes para ingresar a una aplicación, y ejecutar una acción, asumiendo ser el usuario titular.

La idea de asumir que las etapas 1 y 2 pueden ser traspasadas con facilidad, expone el hecho de reforzar la seguridad en la etapa 3. Es aquí donde se observa la dificultad de esta línea, de identificar a un usuario que opera en el sistema con autorización, pues se desconoce característica alguna del usuario real.

7.2. Línea de investigación relacionada con la extracción de información del contexto

Estos sistemas suelen hacer uso de tecnologías complejas y modernas, tales como, interfaces gestuales, sistemas de detección de voz humana, sistemas de reconocimiento de rostro, sistemas que almacenan y utilizan datos biométricos, entre otros, todos ellos pudiendo combinarse, y generar en conjunto un sistema integrado.

Como ventaja puede decirse que, la detección del usuario fraudulento tiende a ser inmediata. Se observó año a año, una mejora en los tiempos de respuesta

en la comparación, entre imágenes, videos, sonidos, etc., aunque todavía se encuentra en proceso de evolución.

Una desventaja es que, los sistemas de esta línea, suelen ser más caros. Para poder operar en relación directa con un cajero automático, computadora de escritorio o dispositivo móvil, esta línea implica que los medios de operación estén equipados con la tecnología requerida, y se debe contar con los complementos necesarios, para que el sistema pueda hacer uso de los mismos.

Si bien, la tecnología utilizada por esta línea está en auge, todavía quedan muchos temas por solucionarse, para que se pueda confiar en resultados exitosos. Por otro lado, sin tener aun definida una base sobre el problema del fraude, como por ejemplo, sin el conocimiento de características de un usuario fraudulento, esta línea es probable que lleve tiempo para poder obtener resultados que resuelvan el problema del FTC.

La salida más rápida para que esta línea pueda realizar aportes a corto plazo, viene dada por el uso de sensores, y puntualmente se habla del GPS. Existen aplicaciones móviles que identifican al usuario fraudulento por ubicaciones de familiaridad, utilizando GPS, aunque no se expusieron modelos computacionales al respecto aun. En estas aplicaciones, cuando un usuario no se encuentra en un lugar familiar, usual o predecible, el sistema toma esto, como un indicador, de que el usuario puede ser fraudulento. Con este aporte, se observa un cierto grado de conocimiento del usuario titular, y sus acciones regulares, en particular, su posicionamiento en tiempos recurrentes. De todos modos, quedan por evaluarse, varias características del usuario real, que no se identifican sólo con GPS.

Mientras esta línea se va incorporando a la tecnología, es recomendable que se combine siempre con sistemas de la línea de seguridad en la red y en el acceso a los datos, y se apoye puntualmente en la seguridad del dispositivo.

Como otra desventaja, finalmente puede decirse que, esta línea implica una implementación compleja y lenta. En contraparte, por los avances evaluados en diferentes trabajos de investigación, puede verse que estas desventajas se pueden corregir, y así, sacar un mayor provecho al uso de la extensa información que se posee actualmente del usuario titular.

7.3. Línea de investigación basada en aprendizaje

Esta línea refiere a sistemas que se retroalimentan con información del usuario, generando reglas de aprendizaje, que se van presentando de manera dinámica, bajo posibles situaciones fraudulentas.

Se puede combinar con la línea de investigación relacionada con la seguridad en la red, y en el acceso a los datos. De esta manera, se protege la conexión de red, para navegar sobre la misma, con un nivel de seguridad aceptable. Así, luego de ingresar el usuario al sistema, el enfoque se basa especialmente en el comportamiento actual del usuario, comparado con los datos y el 'aprendizaje' almacenado al momento. Este aprendizaje viene dado por reglas establecidas o inferidas a partir de datos historiales, y tipos de datos, previamente definidos para un usuario titular. Luego, se generan patrones de aprendizaje, abstrayendo información 'aprendida', en la cual, se identifica que, ciertas características se repiten continuamente, mientras que otras no lo hacen.

Resultan más flexibles que las líneas presentadas con anterioridad, pues, poseen un gran poder de adaptación al cambio. Se presentó un trabajo interesante en [ZAS2011] donde se muestra el estudio de los patrones de comportamiento, llamado 'reconocimiento de actividades'. Además, muestra un pequeño avance, no necesitar de una etapa de entrenamiento del sistema.

Como desventaja, se ha visto en varios trabajos evaluados, en particular, en los modelos computacionales presentados, que es una línea que no aportó grandes cambios y mejoras en un período extenso de tiempo. Aun en las mejores presentaciones de trabajos, se observa una gran cantidad de falsos positivos y falsos negativos. Esto quiere decir que, en varias pruebas realizadas, los resultados muestran que identifican como usuarios fraudulentos a usuarios titulares (falso positivo), y que no identifican a varios usuarios fraudulentos (falso negativo).

Si esta línea, no logra combinarse con otras, como la línea que extrae información del contexto, y mejoran las técnicas para inyectar información falsa a los historiales de datos, es una línea que tiende a desaparecer.

Capítulo 8

Líneas de investigación a futuro

En este capítulo se describen aquellas líneas de investigación de DFTC, que han evolucionado drásticamente, y aquellas líneas, que prometen seguir evolucionando a futuro.

La tecnología más sobresaliente, en los últimos años, vino acompañada por sistemas de interfaces gestuales, sensores de movimiento (entre otros sensores), cámaras de audio, video e imagen. Inclusive en los cajeros automáticos, se espera la llegada de las interfaces gestuales. Estos sistemas prometen, la identificación inmediata del sujeto, por reconocimiento de imagen, voz y/o video, gestos del usuario titular de la TC, etc. Esta es la definición del usuario, por lo que este es, evaluada en el capítulo 5.

Los dispositivos móviles actuales, poseen varios tipos de sensores, que son capaces de reunir información contextual rica, como la ubicación, las firmas de dispositivos inalámbricos, el ruido ambiental, y las fotografías.

En algunos trabajos, se exhorta a la comunidad de seguridad, para volver a diseñar mecanismos de autenticación, para los usuarios de dispositivos móviles. En particular, por el uso de sensores. También se indica que, en lugar de confiar en un modelo simplista, se debe utilizar información contextual, para desarrollar modelos más matizados, que evalúen el nivel de riesgo del entorno actual del usuario. En [FIS2012], por ejemplo, se muestran varios escenarios, que demuestran que la información contextual, se puede utilizar para evaluar los riesgos, y adaptar los mecanismos de autenticación. En este trabajo en particular, se asume que, esta es un área de investigación muy rica, que tiene mucho por explotar. Por último, alude que, se observan futuros trabajos de investigación, para el desarrollo y la evaluación, de los mecanismos de seguridad dinámica, basados en información contextual.

Se espera que gradualmente se incorporen nuevas técnicas, para la identificación física del usuario. Estas pueden aplicarse a través de sistemas que captan, ingresan y comparan datos, extraídos inmediatamente desde el medio utilizado, como, datos del ambiente y del usuario mismo.

Por las conclusiones obtenidas en el capítulo 6, se puede observar que el control del ingreso al sistema no es lo más relevante, sino más bien, el aspecto y comportamiento que posee el usuario que ya ha ingresado al sistema, para cometer un fraude. Se puede ver este caso, como una manera de aguardar por el ingreso del usuario, para una acción segura de DFTC. Porque una vez que el usuario accede al sistema, se pueden utilizar varias técnicas para identificarlo.

Por último, ya existe suficiente información del usuario, inclusive se puede extraer mucho más del ambiente actual, como para corroborar que este usuario es quien dice ser. Aquí es donde se unen dos caminos. El primero, que refiere a la idea de que, se debe conocer mejor a un usuario para poder definirlo; y el segundo, se aplica a la idea de que, extraer información del ambiente, permite conocer mejor al usuario.

Consecuentemente, se considera a la línea de investigación relacionada con la extracción de información del contexto, como la línea candidata, la que tuvo un fuerte impacto, un rápido crecimiento, y un desarrollo constante.

Por todo lo dicho, se puede afirmar que esta línea, será factible se destaque a futuro en la DFTC, entre las líneas estudiadas en este trabajo.

Anexo A

Tecnología Near Field Comunication

La tecnología NFC (Near Field Communication) está pensada para dispositivos móviles y permite una comunicación inalámbrica entre dos dispositivos cercanos, optimizada para comunicación de corto alcance (máximo 20cm).

El uso principal que está impulsando esta tecnología es el pago con dispositivos NFC. A pesar de sus múltiples posibilidades, este aspecto despierta el máximo interés de las empresas, ya que los móviles con NFC prometen sustituir a las actuales tarjetas de crédito.

Google E-wallet

Google Wallet es un sistema de pago móvil creado por Google que permite a sus usuarios almacenar tarjetas de débitos, tarjetas de crédito, tarjetas de fidelidad, y tarjetas de regalo entre otras cosas, en su teléfono móvil.

Esta aplicación utiliza near field communication (NFC) para hacer pagos con un toque del teléfono en cualquier terminal PayPass habilitado al momento de pagar.

Google mostró la aplicación en una conferencia de prensa el 26 de mayo de 2011.

Anexo B

Phishing:

Algunas de las características de los mensajes de correo de phishing son:

- **Uso de nombres de compañías ya existentes.** En lugar de crear desde cero el sitio web de una compañía ficticia, los emisores de correos con intenciones fraudulentas, adoptan la imagen corporativa y funcionalidad del sitio web real, con el fin de confundir aún más al receptor del mensaje.
- **Utilizar el nombre de un empleado real de una empresa como remitente del correo falso**. De esta manera, si el receptor intenta confirmar la veracidad del correo llamando a la compañía, desde esta le podrán confirmar que la persona que dice hablar en nombre de la empresa, trabaja en la misma.
- **Direcciones web con la apariencia correcta**. Tanto los contenidos como la dirección web (URL) son falsos, y se limitan a imitar los contenidos reales. Incluso la información legal y otros enlaces no vitales pueden redirigir a la página web real.
- **Factor miedo.** La ventana de oportunidad de los defraudadores es muy breve, ya que una vez se informa a la compañía de que sus clientes están siendo objeto de este tipo de prácticas, el servidor que aloja al sitio web fraudulento, y sirve para la recogida de información, se cierra en el intervalo de unos pocos días. Por lo tanto, es fundamental para el defraudador, el conseguir una respuesta inmediata por parte del usuario.

  En muchos casos, el mejor incentivo es amenazar con una pérdida, ya sea económica o de la propia cuenta existente, si no se siguen las instrucciones indicadas en el correo recibido, y que usualmente están relacionadas con nuevas medidas de seguridad recomendadas por la entidad.

Ejemplo de uso de mail falso:

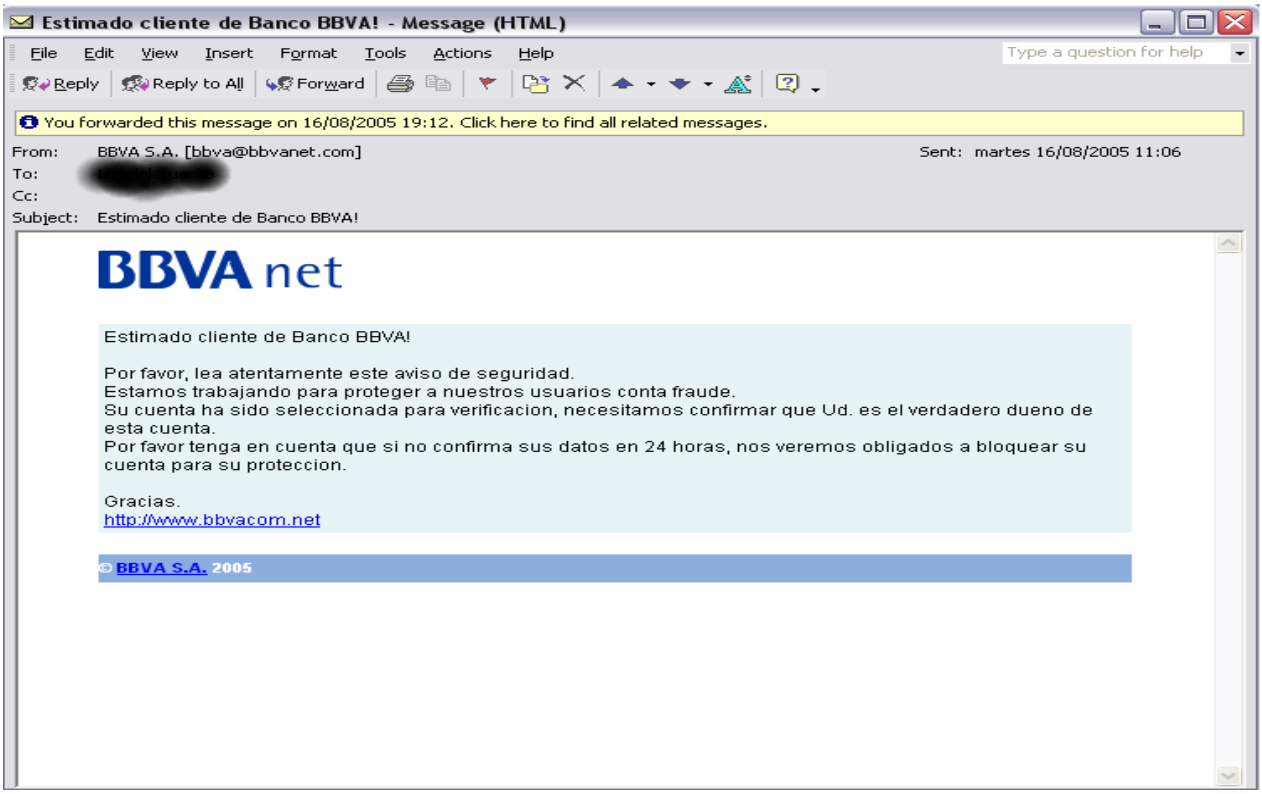

Figura 11. Ataque de phishing.

Actividad en el sitio verdadero:

Figura 12. Ataque de phishing.

Rankings de OWASP 2007, 2010 y 2013:

- **Riesgo de seguridad en las aplicaciones web para el 2007, top 10 [OWA2007]:**

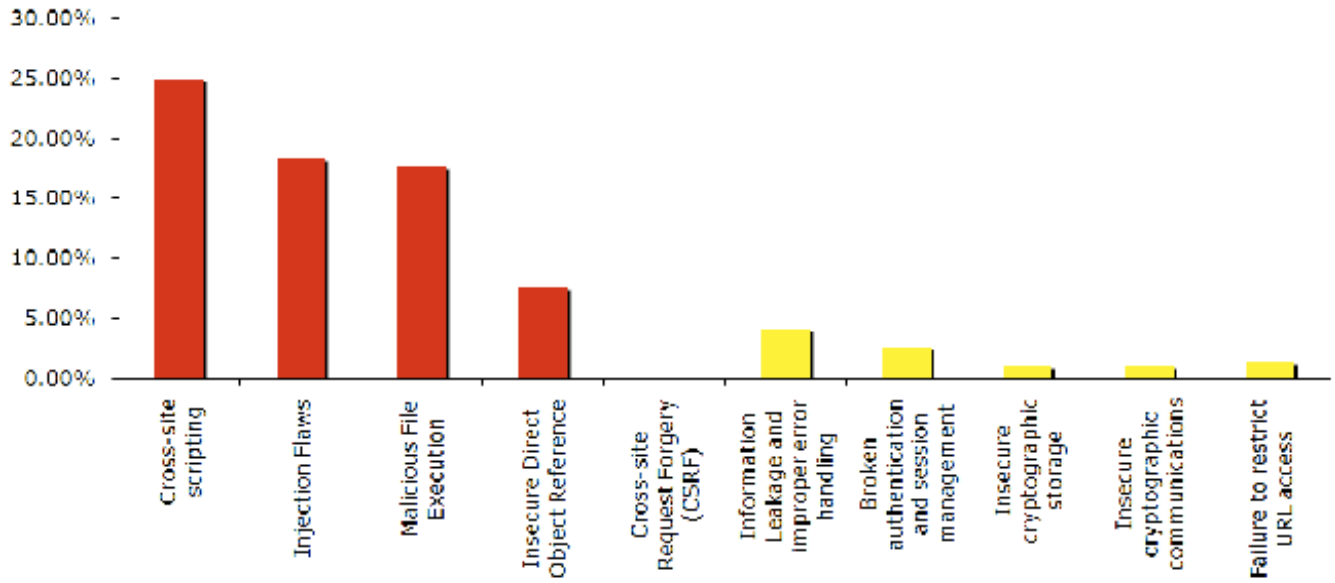

Figura 13. [OWA2007]

- **Riesgo de seguridad en las aplicaciones web para el 2010, top 10 [OWA2010]:**

A1: Injection
A2: Cross-Site Scripting (XSS)
A3: Broken Authentication and Session Management
A4: Insecure Direct Object References
A5: Cross-Site Request Forgery (CSRF)
A6: Security Misconfiguration
A7: Insecure Cryptographic Storage
A8: Failure to Restrict URL Access
A9: Insufficient Transport Layer Protection
A10: Unvalidated Redirects and Forwards

- **Riesgo de seguridad en las aplicaciones web para el 2013, top 10 [OWA2013]:**

A1: Injection
A2: Broken Authentication and Session Management
A3: Cross-Site Scripting (XSS)
A4: Insecure Direct Object References

A5: Security Misconfiguration
A6: Sensitive Data Exposure
A7: Missing Function Level Access Control
A8: Cross-Site Request Forgery (CSRF)
A9: Using Known Vulnerable Components
A10: Unvalidated Redirects and Forwards

Anexo C

4.5. Modelo de Redes Neuronales y Bayesiano

La DFTC usando Redes neuronales y Bayesianas es un sistema automático, por medio del enfoque de aprendizaje automático. Estas dos aproximaciones de aprendizaje de máquina son apropiadas para el razonamiento bajo incertidumbre.

Una red neuronal artificial consiste de un grupo interconectado de neuronas artificiales, y comúnmente, las redes neuronales usadas para la clasificación de patrones es la red con alimentación hacia adelante. Consiste de tres capas llamadas, capas de entrada, oculta y de salida. La secuencia de entrada de transacciones pasa desde la capa de entrada, atraviesa la capa oculta y luego la capa de salida. Esto es conocido como propagación hacia adelante.

El ANN consiste de datos de entrenamiento, los cuales son comparados con la secuencia de transacciones de entrada. La red neuronal inicialmente es entrenada con el comportamiento normal de un titular de la tarjeta. Las transacciones sospechosas son luego propagadas a través de la red neuronal y clasifican las transacciones sospechosas y no sospechosas.

Las redes Bayesianas son también conocidas como Belief Networks (BN), y son un tipo de programación de inteligencia artificial que usa una variedad de métodos, incluyendo algoritmos de aprendizaje de máquina y minería de datos, para crear capas de datos, o concepto (belief). Usando aprendizaje supervisado, las redes Bayesianas, son capaces de procesar datos como se necesite, sin experimentación.

Una BNN es un modelo gráfico que codifica relaciones de probabilidad sobre variables de interés. Cuando es usado en conjunto con técnicas estadísticas, el modelo gráfico tiene muchas ventajas para el modelado de datos.

- Una de ellas es porque el modelo codifica dependencias sobre todas las variables, maneja fácilmente situaciones donde algunos datos de entrada se han perdido.
- La segunda de ellas, es que una BNN puede ser usada para aprender relaciones causales, y por lo tanto, puede ser usado para obtener

conocimiento sobre el dominio de un problema y para predecir las consecuencias de la intervención.

- Otra ventaja, es porque el modelo tiene una semántica tanto causal como probabilística, y esto es una representación ideal para combinar conocimiento previo (que suele estar dado en forma causal) y datos.
- La cuarta ventaja es que, los métodos estadísticos Bayesianos en conjunto con las BNN ofrecen un eficiente acercamiento para evitar el sobreajuste de datos.

  En [EDW2011] puede verse métodos para construir BNN desde el conocimiento previo y se resumen métodos estadísticos Bayesianos para usar datos que mejoran estos modelos. También pueden verse técnicas de aprendizaje con datos incompletos, y de aprendizaje supervisado y no supervisado.

4.6. Modelo adaptativo de detección del fraude

Un método para detectar fraude es, chequear los cambios sospechosos en el comportamiento del usuario. En [HEC1996] se describe el diseño automático de métodos de perfilado de usuario con el propósito detectar el fraude usando una serie de técnicas de minería de datos. Se utiliza un programa de reglas para descubrir indicadores de comportamiento fraudulento a partir de una gran base de datos de transacciones de cliente. Luego, los indicadores son usados para crear un conjunto de monitores, los cuales definen el perfil del comportamiento del cliente legítimo e indican anomalías. Finalmente, las salidas de los monitores son usadas como características en un sistema que aprende a combinar evidencia para generar alarmas de elevada confianza. Este sistema ha sido aplicado para detectar el fraude de clonación celular, basado en una base de datos de registros de llamada.

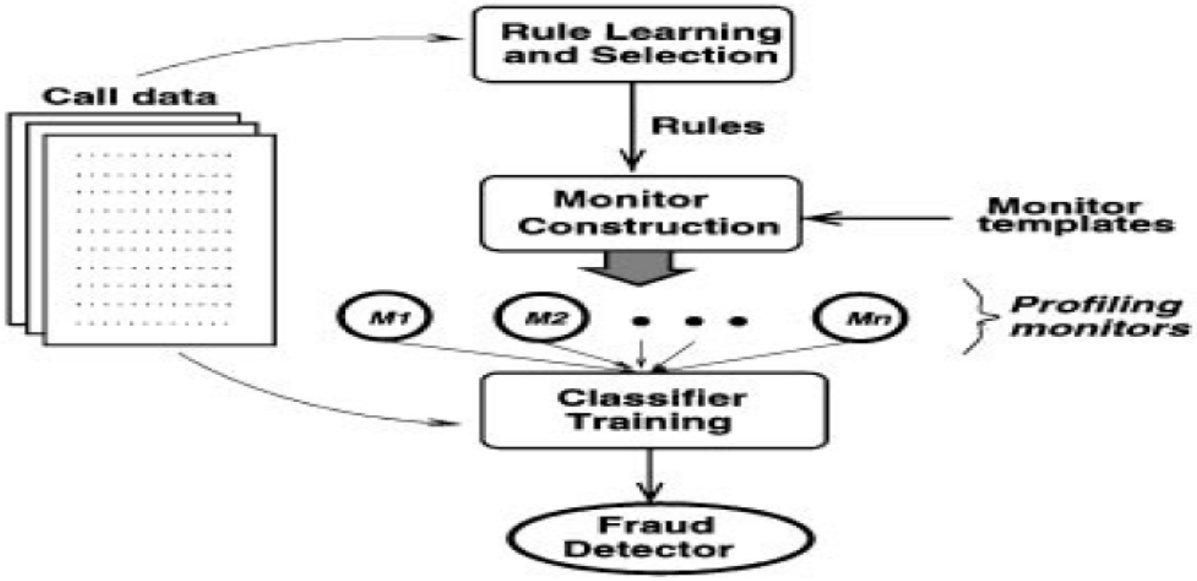

Figura 14. Modelo adaptativo de detección del fraude [HEC1996]

Anexo D

Uso de datos biométricos para la identificación del usuario

Según el centro de difusión de tecnologías de la Universidad Politécnica de Madrid, los sistemas biométricos, también pueden ser pasivos y activos, de acuerdo, al grado de participación del usuario con el sistema:

- Sistemas pasivos. En estos sistemas, el usuario puede no participar en la medida, o bien no saber que es sometido a reconocimiento biométrico. Por ejemplo, el análisis del patrón de voz, del rostro o de los gestos.

- Sistemas activos. En estos sistemas, el usuario participa necesariamente en la medida, y por ello, es consciente de que es analizado biométricamente. Por ejemplo, el análisis de huellas dactilares, de la forma de la mano, del iris, o del patrón de venas de la retina.

La tecnología de los sistemas de identificación biométrica, utiliza características fisiológicas que son estables en los individuos.

Los datos biométricos captados deben:

1. Permanecer constantes en el tiempo para un mismo individuo.
2. Ser diferentes para distintos individuos.
3. Ser de fácil acceso
4. Ser verificables en un tiempo considerable. Por ejemplo, una muestra de ADN es un dato constante en un mismo individuo, y diferente por individuo, pero la extracción de muestras de ADN y el posterior análisis a realizar, no cumplen con las condiciones 3 y 4.

Uso de métodos de ubicación para la identificación del usuario

En [XUL2007] puede observarse una aplicación que realiza la localización a través de GPS, así como de torres de telefonía GSM. A través de los operadores de red usando el área de cobertura de red móvil (MNCA), se detecta automáticamente la ubicación actual del dispositivo móvil. Este mecanismo se produce cuando un usuario con el dispositivo móvil, se registra con el transmisor de la red local a través del cual, los operadores de red

identifican su ubicación e informan al LBAS (Servidor de aplicación basado en ubicación) acerca de la ubicación actual del usuario.

Para registrarse en el sistema, un usuario inicia sesión en el servidor de aplicaciones y entra en sus preferencias seleccionando campos apropiados en formularios del perfil. El servidor de aplicaciones basado en ubicación (LBAS) envía las preferencias a la CPM, que los agrupa en los perfiles de usuario. Si el terminal de cliente tiene un receptor GPS, este envía actualizaciones de ubicación regulares a la CPM. Aunque se centra en los métodos basados en GPS en el que la ubicación del usuario está disponible en el terminal móvil, también se puede utilizar otros métodos de seguimiento de localización, tales como la triangulación basada en la red, que se basa en varios sitios de la célula para determinar la posición de un terminal.

El CPM calcula la dirección del usuario y la velocidad del historial del seguimiento de ubicaciones del usuario, y almacena los últimos valores de contexto dinámico en el perfil del usuario.

Los usuarios móviles que buscan información local sobre el servicio de cercanías, por ejemplo, parques y hoteles cercanos, se conectan al servidor de aplicaciones para solicitar la información. El servidor de aplicaciones reenvía la consulta a la CPM, que consulta al servidor para las direcciones de los servicios Web de restaurante disponibles. El CPM luego envía una solicitud a los servicios Web identificados, especificando un área de servicio para el usuario. Los servicios web buscan sus bases de datos para los recursos apropiados, filtrando aquellos que caen fuera del área especificada, y devuelven una lista XML a la CPM.

Estas formas de localización, pueden ser utilizadas para la detección de un usuario fraudulento, si este se encuentra fuera de ubicaciones de familiaridad.

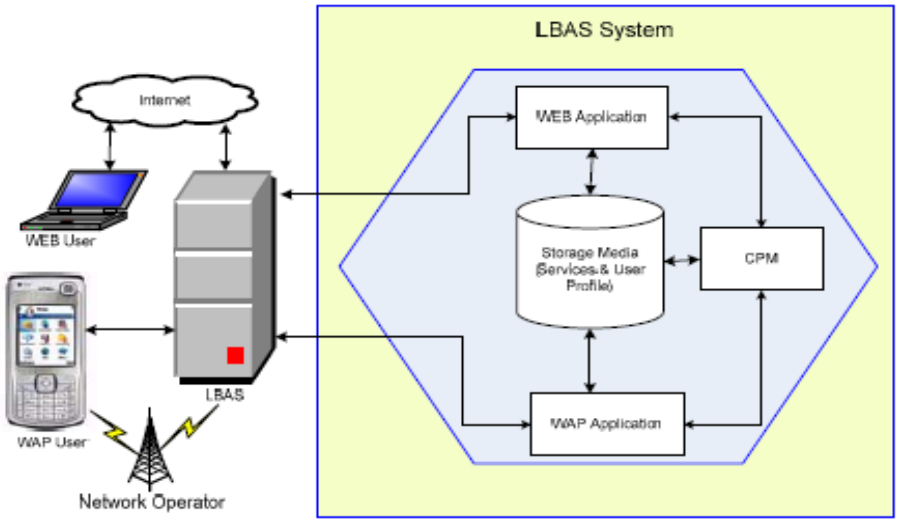

Figura 15. Modelo del método de ubicación del usuario con LBAS System

En [TAJ2011] se puede un interesante sistema WIFI extendido, para la ubicación de dispositivos móviles.

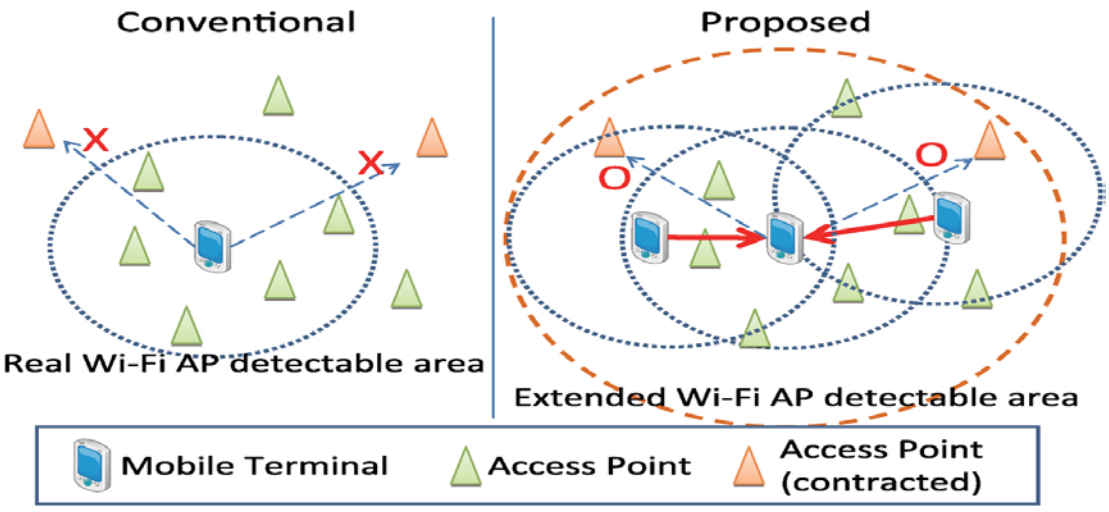

Figura 16. Modelo propuesto de wifi extendido para la ubicación de dispositivos móviles.

El acceso Wi-Fi proporciona una forma muy conveniente, para los terminales móviles, de conectarse a Internet. Sin embargo, la localización de los puntos donde se pueden hacer conexiones no siempre es fácil. En este trabajo, se propone un sistema de anuncios de puntos de acceso Wi-Fi vecinos. Las características del sistema son las siguientes: (1) Una función de punto de acceso personal, opera en terminales móviles que detectan continuamente, y recopilan información sobre puntos de acceso WiFi vecinos (APs), dentro de su rango de detección de radio. (2) La información AP Wi- Fi detectadas y recogidas, así como las ubicaciones y los nombres de red, se comparten entre los terminales móviles en la proximidad entre unos y otros. (3) La información AP Wi-Fi vecinos, puede entonces ser objeto de publicidad a los clientes, a través de sus puntos de acceso personal. En este sistema, un cliente, es un terminal móvil que no tiene privilegios de acceso, para los puntos de acceso cercanos. En tales casos, los clientes no se conectan a Internet a través de los puntos de acceso restringido; estos, obtienen información sobre los puntos de acceso que están disponibles en las cercanías, pero fuera de su rango de detección normal.

Referencias bibliográficas

[ALF2002] Saleh I. Alfuraih y otros. **Using Trusted Email to Prevent Credit Card Frauds in Multimedia Products**. Kluwer Academic Publishers, 2002

[BED2012] Luca Bedogni y otros. **By Train or By Car? Detecting the User's Motion Type through Smartphone Sensors Data**. IEEE, 2012

[BEN2000] Peter J. Bentley y otros. **Fuzzy Darwinian Detection on Credit Card Fraud**. 14th Annual Fall Symposium of the Korean Information Processing Society, 14th October, 2000.

[BER2007] Hal Berghel. **Credit Card Forensics Decoding the magnetic attraction of criminals to swiping.** ACM, 2007.

[BHI1996] Anish Bhimani. **Securing The Commercial Internet**. ACM, June 1996

[CAB2003] Giacomo Cabri y otros. **Location-Dependent Services for Mobile Users**. IEEE, 2003

[CHO2010] Richard Chow y otros. **Authentication in the Clouds: A Framework and its Application to Mobile Users**. ACM, 2010

[EBB2008] P. W. G. Ebben y R. J. Hulsebosch. **Enhancing Face Recognition with Location Information**. The Third International Conference on Availability, Reliability and Security. IEEE, 2008

[EDW2011] S. Benson Edwin y A Annie Porta. **Analysis on Credit Card Fraud Detection Methods**. International Conference on Computer Communication and Electrical Technology. IEEE, 2011.

[FIS2012] Ian Fischer y otros. **Smartphones: Not Smart Enough?**. ACM, 2012

[FUJ2014] www.fujitsu.com Recuperado en 2014

[GUI2008] Tao Guo y Gui Yang Li. **Neural Data Mining for credit card fraud detection**. IEEE, 2008.

[GRA2002] Geraldine Gray. **Virtual Credit Card Processing System**. ACM, 2002

[GRA2007] Jonathan M. Graefe y otros. **Credit Card Transaction Security**. ACM, 2007

[HEC1996] David Heckerman. **Bayesian Networks for Data Mining.** ACM, 1996.

[JUN2010] Junjie Wu y otros. **COG: local decomposition for rare class analysis**. Springer, January 2010.

[JUN2011] Jun Ho An. **Finger gestures based mobile user interface**. IEEE, 2011

[KEL2009] Damian Kellyt y otros. **Computationally Tractable Location Estimation on WiFi Enabled Mobile Phones**. IEEE ,2009

[LEE1996] Chin Teng Lin y C.S. George Lee. **Neural Fuzzy Systems**. Prentice Hall, 1996

[LEE2005] Ronald Leenesk y otros. **Privacy and Identity Management for Everyone**. ACM, 2005.

[MAN2008] Mohammad Mannan y otros. **Localization of Credential Information to Address Increasingly Inevitable Data Breaches**. ACM, 2008

[NAK2009] Tatsuo Nakajima y otros. **UbiPay: Minimizing Transaction Costs with Smart Mobile Payments**. ACM, 2009

[NIT2007] Nitesh Saxena y otros. **Location-Aware and Safer Cards: Enhancing RFID Security and Privacy via Location Sensing**. IEEE, 2007

[OWA2007] www.owasp.org Recuperado en 2007

[OWA2010] www.owasp.org Recuperado en 2010

[OWA2013] www.owasp.org Recuperado en 2013

[PER2012] Rafael Alexandre França de Lima y Adriano César Machado Pereira. **Detecção de fraudes em transações na Web (Fraud Detection in Web Transactions)**. ACM, October, 2012.

[PHU2004] Clifton Phua y otros. **Minority Report in Fraud Detection: Classification of Skewed Data**. ACM, 2004.

[PRE2001] Corinna Cortes y Daryl Pregibon. **Signature-Based Methods for Data Streams**. Kluwer Academic Publishers, 2001.

[QUA2007] Jon T. S. Quah y M. Sriganesh. **Real Time Credit Card Fraud Detection using Computational Intelligence**. International Joint Conference on Neural Networks, IEEE, 2007

[REI1994] Sushmito Ghosh y Douglas L. Reilly Nestor. **Credit Card Fraud Detection with a Neural-Network** . IEEE-Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Sciences, IEEE, 1994

[SCH2008] Saowanee Schou. **Context-based Service Adaptation Platform: Improving the User Experience towards Mobile Location Services**. IEEE, 2008

[SEG2013] www.segu-info.com.ar Recuperado en 2013

[SES2010] Lubor Sesera. **Applying Fundamental Banking Patterns: Stories and Pattern Sequences**. 15th European Conference on Pattern Languages of Programs (EuroPLoP 2010). ACM, 2010

[SHU2010] Shuk Ying Ho y otros. **Users' Adoption of Mobile Services: Preference and Location Personalization.** IEEE, 2010

[SRI2008] Abhinav Srivastava y otros. **Credit Card Fraud Detection Using Hidden Markov Model**. IEEE, 2008.

[STO1999] Salvatore J. Stolfo y otros. **Distributed Data Mining in Credit Card Fraud Detection**. IEEE, December 1999.

[SYE2002] Mubeena Syeda y otros. **Parallel Granular Neural Networks for Fast Credit Card Fraud Detection**. IEEE, 2002.

[TAJ2011] Koji Tajima y otros. **Wi-Fi Access Point Discovery System for Mobile Users**. Eighth International Joint Conference on Computer Science and Software Engineering (JCSSE). IEEE, 2011

[TAK2002] Jun-Ichi Takeuchi y otros. **On-Line Unsupervised Outlier Detection Using Finite Mixtures with Discounting Learning Algorithms**. Springer, 2002.

[UPM2014] www.ceditec.etsit.upm.es Recuperado en 2014

[WAG2011] David Wagner y otros. **A Survey of Mobile Malware in the Wild**. ACM, 2011

[WEI2004] Gary M. Weiss. **Mining with Rarity: A Unifying Framework**. ACM, 2004

[WEN2010] Wenfei Fan y otros. **Dynamic constraints for record matching**. Springer, 2010.

[XUL2007] SS Xulu y otros. **Service supplier infrastructure for location-based m-commerce**. IEEE, 2007.

[YON2003] Yongguo Mei y otros. **Collaborative Intrusion Detection System (CIDS)**: A Framework for Accurate and Efficient IDS. IEEE, 2003

[ZAS2011] Saguna Arkady Zaslavsky y otros. **Towards a Robust Concurrent and Interleaved Activity Recognition of Mobile Users**. 12th IEEE International Conference on Mobile Data Management, IEEE, 2011

Siglas

BNN: Bayesian networks

DF: Detección de fraude

FDS: Fraud Detection System: Sistema de detección de fraude

FTC: Fraude de las tarjetas de crédito

IDS: Sistemas de Detección de Intrusos

PAU: El Problema de la Autenticación del usuario

TC: Tarjetas de crédito