

VALIDACION DE FIRMAS OLOGRAFAS

Tamara D. Blum, Ariel J. A. Sztern, Daniela López De Luise
AI-Group
Facultad de Ingeniería de la Universidad de Palermo
Ciudad de Buenos Aires, Argentina. aigroup@palermo.edu

CONTEXTO

Dentro del ámbito de investigación que se da en el laboratorio de Inteligencia Artificial de la Universidad de Palermo, se viene trabajando sobre un proyecto para la validación de firmas ológrafas o manuscritas.

Este trabajo presenta un prototipo funcional en desarrollo cuyo propósito es facilitar la validación de firmas ológrafas. El sistema realiza la verificación contra una base de datos enriquecida con el agregado de vectores numéricos que sintetizan la imagen de la firma capturada según cierta red neuronal. Además se presenta el diseño del esquema de seguridad y se justifican las decisiones arquitectónicas desde un punto de vista teórico.

RESUMEN

Existen, al día de hoy, una cantidad de alternativas en lo que hace a la validación de individuos. La mayoría de éstas presentan una serie de inconvenientes que las hace difícilmente masivas y/o suficientemente seguras. El presente trabajo profundiza acerca de los conceptos involucrados en dichas alternativas, y presenta la implementación práctica de una Solución Informática para validación de firmas manuscritas, denominada proyecto PI-1. El prototipo obtenido responde a una arquitectura original e introduce variantes que han surgido como producto de la investigación realizada en la combinación y optimización de ciertas técnicas. Se han concentrado los esfuerzos en obtener un procesamiento de tipo off-line de las firmas (denominado estático) a partir de imágenes capturadas desde un dispositivo celular móvil.

Palabras clave: Validación de firmas, Firmas manuscritas, Firmas ológrafas, Biometría, Criptografía

1. INTRODUCCION

El prototipo PI-1 permite calcular el grado de cercanía que tiene una firma capturada respecto de otras previamente registradas para un determinado firmante, a la vez que valida otros datos identificatorios. Ello ayudará a validar la identidad de un individuo, preservando el sentido legal que tiene la firma manuscrita.

Algunas de las razones que motivan esta investigación son:

- Los dispositivos biométricos que existen actualmente deben ser complementados para cumplir con los requisitos esperados en cuanto a la seguridad o son métodos invasivos y de difícil aplicación.
- Aún existiendo la Ley de Firma Digital [1][2][3][4][5][6][7], la firma manuscrita es un requisito ineludible en gran cantidad de trámites, como transacciones inmobiliarias, cesiones de poder, autorizaciones o documentos jurídicos [8][12].
- Es un método en el que se puede obtener un bajo índice de error.
- No se requiere que el firmante sepa leer y escribir, dado que las firmas son simplemente garabatos.
- Es una forma natural de identificarse, ya que es costumbre complementar ciertos trámites con la rúbrica personal en la documentación.
- No se ve afectada por implicancias externas como puede ser el tipo de trabajo que realice una persona, como si sucede con la huella dactilar [9], donde puede deteriorarse el dibujo de la misma.
- No requiere de costoso equipamiento.

Entre las características establecidas para el

prototipo se han definido las siguientes:

- Debe ser portable.
- Capacidad de ser usado en forma masiva.
- Debe ser accesible vía Internet.
- No debe tener costos adicionales por licenciamientos de terceras partes, como motores de bases de datos, sistemas operativos, etc.
- Debe poder ser ejecutable desde dispositivos móviles, como también en servidores o computadoras personales.
- No puede depender de la existencia de hardware de alto costo como una limitación funcional.
- Debe proveer un nivel de seguridad aceptable para el entorno de uso, ya que los datos que se manejan pueden llegar a ser sensibles.

El prototipo consta de una base de datos que debe ser alimentada con la información de los firmantes y una muestra de firmas (se recomienda no menos de 5 firmas por firmante para garantizar la calidad de los resultados). Dichas muestras son las que permiten extraer las características de las firmas de cada firmante por una red MultiLayer Perceptron (MLP), que se usará luego para poder validar las futuras muestras y devolver un grado de coincidencia respecto de las muestras existentes.

Si bien la propuesta originalmente contempla la captura de firmas utilizando como fuente de información diferentes alternativas y dispositivos, la implementación realizada se centró en el uso de dispositivos celulares móviles. Sin embargo debe destacarse que la forma en que se provee la entrada influye en la calidad de resultados que se obtienen.

2. PROTOTIPO PI-1

A continuación, se describen brevemente los puntos más importantes que se han tomado en cuenta para el prototipo.

Arquitectura general

Como se ha mencionado anteriormente, la solución que se plantea permitiría que se incorporen y validen firmas a partir de

diferentes formatos y dispositivos: dispositivo de lectura on-line, archivo de firmas (escaneado o fotocopiado), o Webcams. El desarrollo admite ser invocado a través de Internet como un servicio Web.

Las posibilidades que brindan los dispositivos de lectura on-line son mayores, motivo por el cual se puede ofrecer un mejor resultado mediante este tipo de procesamiento. No obstante, mediante la lectura off-line también se puede llegar a obtener resultados con un bajo índice de error.

El esquema de funcionamiento se muestra en la figura 1.

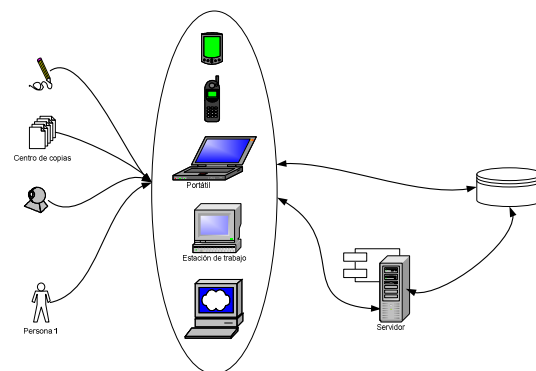


Fig. 1. Arquitectura general del sistema de validación.

Previo a todo proceso de validación es imprescindible que se cargue una serie de muestras de cada individuo contra las cuales validar. A fin de contemplar los cambios naturales que se van dando en las firmas de los individuos es recomendable que las mismas sean actualizadas a lo largo de los años.

PI-1 analiza la firma y la registra en la base de datos de firmas; o bien sólo la valida dependiendo de los requerimientos eventuales.

Base de Datos

La base de datos se almacena en formato XML (Extensible Markup Language). Esto permite tener un mejor control a nivel seguridad, y la hace totalmente portable ya que no requiere motores específicos para el acceso.

Ya se ha mencionado anteriormente que dentro de los objetivos de la propuesta se encuentra el obtener un producto capaz de funcionar en diferentes dispositivos y/o plataformas. Por lo tanto, no se puede dejar de tener en cuenta que

motores de bases de datos existentes podrían llegar a ser un inconveniente en este aspecto. La situación se agrava aún más si requiere de costos adicionales en concepto de licencia. Es por ese motivo que se ha decidido trabajar con una base de datos XML.

Adicionalmente, este tipo de bases de datos suponen un acceso a la información más performante que las bases de datos relacionales, dado que guardan los documentos enteros físicamente juntos o utilizan punteros físicos en lugar de lógicos entre las partes de los documentos [10].

Las bases de datos XML son del tipo árbol, donde los nodos son los atributos y cada una de las subdivisiones son las raíces jerárquicas.

Existen actualmente diferentes alternativas para acceder a la información desde distintos lenguajes de programación. Y no necesariamente hay una forma que sea la más conveniente. Los requisitos para validar no son los mismos que para la registración.

Se optó por desarrollar una interfaz de acceso a los datos e implementarla utilizando SAX. Esto permite intercambiar la forma de acceder a los datos en un futuro, si se encontrara una mejor alternativa, con sólo implementar en una nueva clase la misma interfaz.

Seguridad

Los datos se guardan encriptados para dificultar la interpretación de los mismos e impedir el “robo” de información.

Sin embargo, teniendo en cuenta que pueden producirse ataques de distintos tipos, dentro del presente esquema, se trabaja la seguridad en todos los niveles de la arquitectura. Esta se brinda a partir de:

- Encriptación GSM (Global System for Mobile Communications)
- Páginas Web Seguras
- Encriptación en Servicios Web
- Seguridad en Threads

El procesamiento de los datos contempla altos niveles de seguridad, ya que se utilizan los propios de las conexiones Web o móviles y también métodos adicionales para dar mayor seguridad al proceso y prevenir robo de información en el momento del procesamiento

de los datos.

Interfaces gráficas

Las interfaces gráficas son simples e intuitivas, tratando que los datos sean visibles en su conjunto dentro de la misma ventana gráfica. Ello agiliza y facilita el uso de la solución.

A continuación se grafican las interfaces principales del sistema.

En las figuras 2.1, 2.2 y 2.3 se muestran las pantallas de registración de un nuevo firmante, con los datos básicos necesarios. El primer caso grafica como sería desde un teléfono celular y el segundo desde una computadora personal.

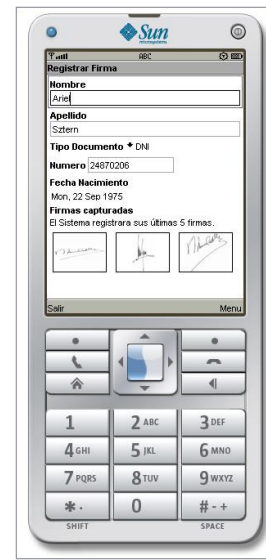


Fig. 2.1. Formulario para captura de datos del firmante y muestras de firmas desde teléfono celular.

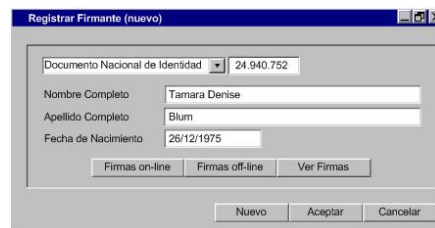


Fig. 2.2. Formulario para registrar un nuevo firmante desde una computadora personal.

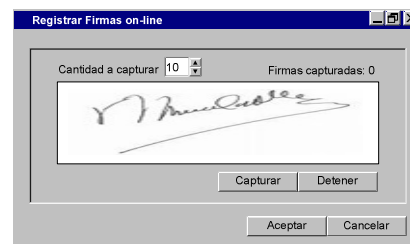


Fig. 2.3. Formulario para registrar firmas utilizando un dispositivo de captura (on-line).

Procesamiento de imágenes

El análisis de las características y patrones comprendidos en las firmas ológrafas, propone un desafío complejo, dado que el mismo, varía dependiendo de la calidad de la muestra obtenida.

Teniendo en cuenta que los dispositivos utilizados para la captura de firmas, son variados y con características diferentes (celulares, PDAs, archivos, etc.), es necesario realizar un proceso que permita normalizar y eliminar aspectos que dificulten el procesamiento de las imágenes en la red neuronal.

Se han analizado distintas alternativas que tienden a mejorar la calidad de los registros de firmas obtenidos, enfocándose principalmente en los siguientes aspectos:

- **Ruido:** el ruido comprende cualquier alteración producida antes o durante la captura, que incorpore trazos, marcas, manchas o sombras, externas a la firma. Esta deformación de la firma, puede generar problemas en la identificación de sus características.
- **Centrado:** teniendo en cuenta que el método para captura de las firmas no se encuentra estandarizado, es altamente probable que las imágenes obtenidas no se encuentren centradas. Esto distorsiona el reconocimiento de los patrones, entre firmas de una misma persona.
- **Tamaño:** las muestras obtenidas, al no ser uniformes, pueden responder a distintas escalas. Esto dificulta no sólo la comprensión de los patrones de la firma, sino también ocasiona un mayor tiempo de procesamiento por parte de la red neuronal, ya que las imágenes contienen mayor información. Para ello, es fundamental reducir el tamaño de la muestra, generando condiciones similares para el procesamiento de todas las firmas.
- **Color:** se ha considerado para el presente trabajo, la estandarización de los colores correspondientes a la firma, con el fin de

normalizarla y reducir su volumen. Por ello, se determinó utilizar imágenes en blanco y negro, lo que permite simplificar la obtención del vector de características y agilizar el procesamiento.

Identificación de patrones

El punto más crítico de la investigación, se ha focalizado en la identificación de los patrones característicos de una firma ológrafa. Estos patrones son los que identifican aquellas particularidades que vuelven única a la firma. Se han estudiado distintas aproximaciones utilizadas en campos similares:

- **Reconocimiento de rostros:** en este punto se ha analizado la propuesta introducida por Sirovich y M. Kirby, conocida como Eigenfaces [11]. A diferencia de un rostro, la identificación de firmas, es una tarea que requiere mayor exactitud, ya que al validar una firma, se debe desestimar aquellas que correspondan a falsificaciones. En el caso de los rostros, si dos personas poseen facciones muy similares, es probable se deseen obtener ambos resultados como válidos.
- **Redes neuronales de Clasificación:** estas redes responden a patrones más específicos, pudiendo clasificar una determinada entrada, en un conjunto de salidas ya conocidas. En este caso, el resultado obtenido no es un vector de características, sino una clase específica, previamente establecida durante el entrenamiento de la red. La complejidad presentada por las firmas, hace que la identificación concreta de una salida, puede desembocar en falso positivo, al catalogar como válida una falsificación, por poseer mayor afinidad con la firma real que con el resto del universo de firmas.
- **Redes neuronales de Retropropagación:** hemos considerado este tipo de redes, como la aproximación más concreta sobre la problemática planteada, dado que es capaz de ajustar sus pesos en forma automática a partir de los valores obtenidos. Esto contribuye a mejorar la

comprensión de los rasgos que componen la firma, obteniéndose como resultado un vector que representa los patrones significativos de la firma procesada.

3. RESULTADOS Y TRABAJO A FUTURO

Al momento se han desarrollado los prototipos para captura de imágenes y el almacenamiento de las firmas y demás información acerca del firmante en la Base de Datos.

Se ha implementado la seguridad en los procesos de comunicación y resguardo. Se está trabajando en las alternativas de procesamiento con redes neuronales.

Queda como trabajo futuro finalizar el estudio de las redes neuronales y las alternativas de implementación. También se realizará un estudio de eficiencia y confiabilidad.

Asimismo, se realizará el análisis estadístico de confiabilidad en validaciones on-line (dinámico) de firmas manuscritas.

Respecto al diseño de la base de datos, queda pendiente terminar de definir si la mejor alternativa es que los datos se guarden en su conjunto dentro del mismo archivo o guardarlos estilo “cabecera-detalle”.

Las ventajas de guardar el conjunto de datos en diferentes archivos es que ante el hecho de corromperse alguno de ellos, se puede rearmar la base de datos con sólo tomar nuevamente las muestras que corresponden al firmante que tiene sus firmas corruptas. Sin embargo, el resguardo de la información puede verse afectado si el archivo que se corrompe es el que actúa de “cabecera”. El mismo grado de dificultad existe ante la corrupción del archivo que contendría a la base de datos completa,

Aunque el almacenamiento de la información utilizando varios archivos es ventajoso por su disposición, este sistema podría traer como consecuencia alguna desventaja en la seguridad, ya que permitiría un rastreo más simple del contenido y las muestras de las firmas por parte de usuarios locales del sistema.

4. FORMACION DE RECURSOS HUMANOS

En el marco de la red de colaboración PI, dedicada a la implementación de sistemas inteligentes, suman sus aportes profesionales de las siguientes instituciones:

- Por la Universidad de Palermo intervienen 1 doctora y 2 licenciados en informática.
- Por la UADER intervienen 1 doctor y 2 doctorandos en roles de investigación.
- Por la UNLP interviene 1 doctor con rol de investigador.
- Por la UTN Facultad Regional Concepción del Uruguay intervienen 3 graduados como investigadores.

5. BIBLIOGRAFIA

- [1] Información consultada en <http://www.saij.jus.gov.ar>, Sistema Argentino de Informática Jurídica - Ministerio de Justicia, Seguridad y Derechos Humanos
- [2] Ley 25.506, Ley de Firma Digital, sancionada el 14 de Noviembre de 2001. <http://www.saij.jus.gov.ar>
- [3] La Ley Provincial N° 12.491, del 24 de Noviembre de 2005 dispone de un plazo de hasta 5 (cinco) años para la implementación de la tecnología necesaria para la Firma Digital en el ámbito normativo.
- [4] Ley Provincial N° 633, del 6 de Julio de 2004.
- [5] Ley Provincial N° 2.073, del 9 de Octubre de 2003.
- [6] Ley Provincial N° 5.366, del 5 de Julio de 2005.
- [7] Ley Provincial N° 13.666, del 12 de Abril de 2007.
- [8] Trevathan, J. and McCabe, A. Remote Handwritten Signature Authentication. In Proceedings of the 2nd International Conference on eBusiness and Telecommunication Networks (ICETE '05), pp 335--339, 2005. <http://citeseer.ist.psu.edu/trevathan05remote.html>.
- [9] Gupta, J and McCabe, A. A Review of Dynamic Handwritten Signature Verification. James Cook University, Australia (1997). <http://citeseer.ist.psu.edu/gupta97review.html>
- [10] Ronald Bourret, “XML and Databases”, Septiembre 2005, <http://www.rpbouret.com/xml/XMLAndDatabases.htm>.
- [11] L. Sirovich and M. Kirby, “Low-dimensional procedure for the characterization of human faces”, Journal of Optical Society of America. 1987
- [12] G. Rodríguez, “De la firma autógrafa a la firma digital”, Facultad de ciencias jurídicas y políticas, Universidad de Zulia [Disertación sobre firma ológrafa y digital; y la implementación de ésta última como medio de autenticación y confirmación para documentos electrónicos].