Seguridad en la Nube: Almacenamiento de Imágenes Médicas Y Watermarking

Silvia Edith Arias[1], Laura Mónica Vargas[2,3], Alejandra Di Gionantonio[1], Daniel Arch[1], Diego Serrano[1], Martín Navarro Mugas[1], Nicolás Hernandez[1], Paula Sosa[1], Ezequiel Ambrogio[1]

s_autn@hotmail.com, laura.monica.vargas, ing.alejandradg, diegojserrano, mnavarromugas, damiannicolas05, sosa.pau, ezequielambrogio@gmail.com, daniel.arch@pjn.gov.ar

[1]Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de Información, Facultad Regional Córdoba, Universidad Tecnológica Nacional

[2] Laboratorio de Redes y Comunicaciones de Datos, Departamento de Computación, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba

[3] Laboratorio de Procesamiento de Señales, Departamento de Matemática, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba

I. Resumen

En los últimos años, se desarrollaron muchas aplicaciones acompañando el rápido avance de las telecomunicaciones. Una de ellas es la telemedicina por medio de la cual los médicos pueden transferir y compartir los datos digitales de los pacientes en forma remota para determinar un diagnóstico definitivo. Actualmente, la información médica que se almacenaba en el centro de salud se lleva a la nube. Por lo tanto, es esencial proteger los datos médicos intercambiados, especialmente cuando se utiliza una plataforma de Cloud Computing donde la seguridad es un problema importante. Hay que garantizar que las imágenes médicas se puedan compartir en forma segura preservándolas de cualquier intento de distorsión. Como así también proporcionar privacidad en las cadenas de datos de los Registros Electrónicos de los Pacientes o Electronic Health Records (EHR).

En este trabajo, nos dedicamos a explorar y obtener conocimientos teóricos sobre marcos de trabajo de plataformas de Health Cloud Computing que permitan alojar imágenes médicas con inserción de Marcas de Agua en los EHRs. La finalidad de este trabajo de campo es seleccionar y probar plataformas de Health Cloud Computing. Posteriormente elaborar una comparación cuantitativa y cualitativa de sus características principales, y recomendar cuáles son las plataformas más seguras y adecuadas.

Palabras clave: seguridad informática, watermarking, imágenes digitales, cloud computing, telemedicina.

II. Contexto

El presente trabajo se realiza en el Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de Información de la Facultad Regional Córdoba de la Universidad Tecnológica Nacional. En el marco del Proyecto "Análisis comparativo entre Plataformas de Cloud Computing, para el caso de almacenamiento de imágenes médicas con marcas de agua" acreditado y financiado por la Secretaría de Ciencia y Técnica de Código: CCUTNCO0004961. El cual se lleva a cabo en el Laboratorio de Investigación de Software de la Facultad Regional Córdoba de la Universidad Tecnológica Nacional (Argentina).

La temática de watermarking ha sido presentada por los cuatro primeros autores en los Proyectos homologados por SeCyT-UTN PID, "Marcas de Agua múltiples en imágenes digitales fijas para autenticación y detección de adulteraciones". Código SCyT – UTN1166, 2010-2011, Resolución 26/10, 2010 SCyT del Rectorado de UTN y "Marcas de Agua Seguras en Imágenes para identificación del propietario". Proyecto ID

promocional. Código SCyT- UTN EIPRCO753, 2008-2009, Resolución 75/08 SCyT del Rectorado UTN.

La segunda y la tercera autora han publicado en la Revista de la FCEFyN de la UNC un artículo de difusión de marcas de agua. "Marcas de Agua: una Contribución a la Seguridad de Archivos Digitales". Revista de la Facultad de Ciencias Exactas, Físicas y Naturales de la UNC. ISSN 2362-2539 (Versión electrónica). Año 3 – Nº 1 (2016).

III. INTRODUCCIÓN

En la etapa actual, la investigación se centra en el análisis del problema de seguridad de la información contenida en las imágenes médicas cuando estas son alojadas en un framework de Cloud Computing.

Cloud Computing es un mecanismo que creció en los últimos años, basado en la Web que permite escalar y virtualizar recursos de TI que son proporcionados como servicios a través de la red. Características inherentes y esenciales que deben ser provistas por las aplicaciones de cloud computing son: servicio bajo demanda, acceso ubicuo, escalabilidad, elasticidad, independiza al usuario del mantenimiento y pago por uso, siendo la seguridad todavía un desafío [1] [2].

Por otra parte, en los últimos años se ha logrado un progreso significativo en el uso de tecnologías de comunicación para almacenar y distribuir datos médicos bajo formatos digitales [3].

El uso de las mismas no es seguro cuando los datos, médicos o de otro origen, circulan libremente por redes abiertas como Internet, expuestos a que los mismos sean alterados o mal utilizados. Esto sucede especialmente cuando se emplean servicios de teleconsulta o telediagnóstico, que se están difundiendo por todo el mundo con el aporte de las Tecnologías de Informática y Comunicaciones. Un esquema de implementación se muestra en la Figura 1.

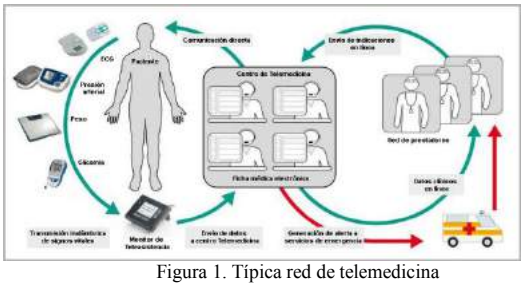

Figura 1. Típica red de telemedicina

La necesidad de tomar medidas de seguridad se incrementó con el almacenamiento de información médica en la nube.

Almacenar en la nube es una buena alternativa ya que permite a los centros médicos desentenderse del hardware y software usado, pero conlleva un mayor riesgo de violación de autenticidad e integridad en los registros del paciente. La confidencialidad de los datos es una necesidad ética en el campo de la salud. La tecnología de encriptación clásica es una herramienta importante que puede y debe ser utilizada para proteger los datos transmitidos en redes de computadoras [4], pero no es suficiente para solucionar todos los problemas de protección de datos digitales.

Además se debe considerar que, en la actualidad, los sistemas de información no son centralizados, sino distribuidos y que, por lo tanto, el control también debe estar distribuido. La protección primitiva que controlaba el acceso, mediante claves, hoy no es suficiente.

El watermarking o marcado de productos multimedia, imágenes, videos, audio, gráficos, etc., se empezó a desarrollar desde la década del 90 como forma de protección de propiedad intelectual [5]. Consiste en embeber bits en el archivo, sea imagen, video o audio, de forma visible (audible) o invisible (no audible). Estos bits extra constituyen la marca y en las primeras implementaciones permitían identificar al propietario, utilizándose posteriormente para alcanzar otros propósitos como detección de adulteraciones, aseguramiento de integridad e incorporación de metadatos. Esta técnica se utilizó posteriormente para conseguir otros propósitos como detección de adulteraciones, aseguramiento de integridad e incorporación de metadatos. Así, en estos días, la marca de agua aparece como un medio eficiente para asegurar integridad y verificar autenticidad. Actualmente, se está adaptando esta técnica para imágenes médicas, siendo de

particular interés el embebido de los datos médicos del paciente en sus imágenes personales [6]. Resultados experimentales, mostraron que la aplicación técnica de marcas de agua a este tipo de imágenes las tornaron más resistentes a varios tipos de ataques [7]. Las amenazas, como la destrucción de sistemas de software y violación en los accesos, están surgiendo con frecuencia en la plataforma de la nube, por lo que se hace absolutamente necesario tomar medidas para contrarrestarlas [8] [9].

Las imágenes médicas son almacenadas por los siguientes dos propósitos:

- Diagnóstico
- Base de Datos (almacenamiento a largo plazo)

Las imágenes deben ser guardadas perfectamente sin ninguna pérdida de información antes de que el médico haga su diagnóstico. Deben ser almacenadas sin compresión o comprimidas mediante un algoritmo que no pierda información.

Al comienzo de los años 80 apareció el PACS (Picture Archiving and Communication System/ Sistema de Comunicación y Almacenamiento de Imágenes). Inicialmente, los PACS se desarrollaron para cubrir necesidades específicas, tales como adquisición de datos y visión de estos en estaciones de trabajo de poca capacidad [10]. Rápidamente se extendieron en las instituciones médicas, pero presentaban problemas de interoperabilidad ya que eran desarrollados en forma independiente por distintos proveedores que seguían sus propias reglas. Surgió entonces la Norma DICOM (Digital Imaging and Communications in Medicine), desarrollada por el ACR (American College of Radiology) en conjunto con NEMA (National Electrical Manufacturers Association) [11]. Tras varios intentos se aprobó en 1993 y sufre una actualización constante. Esta norma define el acceso a la web, la estructura de intercambio, las capas de comunicación de datos y los comandos para manejo de imágenes médicas que deben respetar todos los fabricantes para obtener interoperabilidad. Su aceptación hizo que se "dicomizaran" los PACs. Actualmente los fabricantes de equipos para imágenes médicas siguiendo indicaciones de la norma los acompañan de un CS (Conformance Statement) que asegura que cumplen con la misma. Esta norma permite el acceso remoto a archivos en formato DICOM (extensión dcm) utilizando los ya clásicos protocolos TCP/IP, emplean el Protocolo HTTP (Hypertext Transfer Protocol) o HTTPS (Hypertext Transfer Protocol Secure). Si bien aseguró interoperabilidad entre los distintos sistemas y demostró cierta flexibilidad en entornos que manejan imágenes médicas, no hizo un aporte significativo a la seguridad ni al acceso de datos por fuera de instituciones médicas

En 1996, se dictó en EEUU, la HIPAA (Health Insurance Portability and Accountability Act) que indica qué requisitos se deben cumplir para las transacciones de datos de salud con el objetivo de que los datos médicos se almacenen y se puedan recuperar a largo plazo, evitando abusos y fraude. Microsoft en 2007 y Google en 2008 ofrecieron portales Health a los usuarios que querían que sus EHRs estuvieran disponibles para sus servicios de salud y para ellos mismos. En 2010, IBM y Aetna en conjunto anunciaron un nuevo uso de la plataforma de cloud computing de IBM diseñada para ayudar a los profesionales de la salud a acceder rápidamente a la información del paciente: registros médicos, recetas, y datos de laboratorio recolectados de múltiples fuentes para crear un registro detallado del mismo. Se estima que en el año 2020 el 80% de los datos se habrá mudado a la nube. El uso de estas plataformas y otras permite a los centros médicos desentenderse de problemas técnicos (actualización y mantenimiento de software y hardware), económicos y legales relacionados al manejo de datos lo que le conviene más allá de los riesgos que corre. Entre los inconvenientes se encuentra su latencia, la dificultad para tener el servicio disponible todo el tiempo, y la seguridad [12]. Se debe tener especialmente en cuenta que los datos almacenados en la nube son vulnerables a ataques internos. La identidad y ubicación de intermediarios y de los proveedores de servicio está disimulada, oculta, por la nube.

La importancia del watermarking en imágenes médicas ya se destacaba dos décadas atrás [13], y sigue considerándose en la actualidad [14]. Es de particular interés que se embeban los datos del

paciente en sus imágenes médicas personales. Las amenazas, como la destrucción de sistemas de software y violación en los accesos, están surgiendo con frecuencia en la plataforma de la nube, por lo que se hace absolutamente necesario tomar medidas para contrarrestarlas.

Se recomienda en telemedicina, la combinación de watermarking con técnicas de criptografía clásicas [13-17]. La encriptación puede impedir problemas en los nodos intermedios, pero no en los puntos finales que deben poder desencriptar los datos y si el proveedor del servicio confía, a su vez, en otros proveedores entonces los datos del usuario pueden ser leídos por muchas entidades en la nube. Lo que se precisa para incrementar la confianza en la nube siguiendo esta línea de razonamiento es algún mecanismo que pueda detectar y castigar cualquier problema relativo a la confidencialidad. El usuario final debe confiar en la entidad que administra los EHR (sea Azure, Amazon, Google Health, Microsoft Health Vault u otro proveedor de nube). El administrador de EHR debe tener los medios para detectar y castigar las violaciones a la confidencialidad que se hubieran producido. Las nuevas técnicas deben tratar especialmente de mitigar los riesgos de que la información sufra ataques internos en la nube [18].

La propuesta actual de investigación consiste en realizar una comparación entre diversas plataformas de cloud computing para alojar imágenes médicas marcadas mediante un mecanismo que combine watermarking y encriptación para lograr seguridad, integridad y autenticidad de datos médicos. Se propone utilizar un sistema de archivos distribuidos Hadoop, usar MapReduce, modelo de programación para el manejo de grandes bases de datos [19].

IV. Objetivos y Líneas de Investigación y Desarrollo

El proyecto se inscribe dentro de los lineamientos de investigación en Seguridad Informática.

Para el desarrollo de este trabajo de investigación se aplicará el método empírico-analítico, que se basa en la experimentación y en la lógica empírica, junto a la observación de plataformas de Nubes para alojar imágenes médicas con marcas de agua.

El objetivo de este proyecto de investigación es analizar el estado del arte de cloud computing para servicio de almacenamiento de imágenes médicas con marcas de agua y difundir los resultados obtenidos para realimentar el proceso de desarrollo de los algoritmos de watermarking.

V. Materiales y Métodos

En nuestra investigación proponemos hacer frente al problema de la seguridad de los datos contenidos en imágenes médicas alojadas en Cloud Computing, utilizando la técnica de marca de agua en el EHR. Luego se envía la imagen con marca de agua al proveedor de la Nube. En particular se trabajará con imágenes médicas en formato dicom y con marcas reversibles indetectables.

Es común que una imagen médica sea diagnosticada antes de que la misma sea almacenada en un almacenamiento a largo plazo, de este modo la parte significativa de la imagen, conocida como ROI (Region of Interest), es determinada en ese momento, con lo que el embebido de información extra se puede hacer fuera de esta zona.

VI. Formación de Recursos Humanos

El grupo está compuesto por una Directora, cuatro profesores investigadores de apoyo, tres ingenieros aspirantes a incorporarse a la carrera de investigador y un estudiante investigador de la carrera de Ingeniería en Sistemas de Información. Este proyecto contribuirá a la formación y crecimiento de la carrera de investigador de los integrantes del mismo.

Además existe la colaboración de una docente investigadora de la FCEFyN-UNC.

El desarrolla tareas de investigación en el Laboratorio de Investigación de Software (Lis).

Se dirigirán trabajos finales sobre la temática abiertos a estudiantes de Ingeniería en Sistemas de Información.

Referencias

[1] Youssef *et al.* “Toward a Unified Ontology of Cloud Computing”. Grid Computing Environment Workshop, IEEE, 2008.

[2] Jadeja and Modi. “Cloud Computing – Concept, Architecture and Challenges”.

International Conference on Computing, Electronics and Electrical Technologies, IEEE, 2012.
[4] Stallings. "Cryptography and Network Security", Ed. Prentice Hall, 4th Ed., 2005.
[5]Cox, Miller y Bloom - "Digital Watermarking" - Morgan Kauffmann, 2002.
[6] Coatrieux et al- "Relevance of Watermarking in Medical Imaging"- Proceedings of the IEEE EMBS Conf. on Information Technology Applications in Biomedicine, Arlington, USA, Nov. 2000, pp 250-255.
[7] Fatma E.-Z. A. Elgamal, Noha A. Hikal, F.E.Z. Abou-Chadi– *"Secure Medical Images Sharing over Cloud Computing environment"*, Information Technology dept. Faculty of Computers and Information Sciences, Mansoura University Mansoura, Egypt . 2013.
[8] Junning Fu et al- *"A Watermark-aware Trusted Running Environment for Software Clouds"*, School of Software Tsinghua University Beijing 100084, China, 2009.
[9] Yu-Chao Liu et al- *"A Method for Trust Management in Cloud Computing: Data Coloring by Cloud Watermarking"*, Department of Computer Science and Technology, Tsinghua University, Beijing, 2011.
[10] Bharath. "Introductory Medical Imaging". Ed. John Enderle, University of Connecticut, 2009.
[11]Pianykh. "Digital Imaging and Communications in Medicine (DICOM). A Practical Introduction and Survival Guide". 1st Ed., Ed Springer, 2008.
[12] Zhang and Liu. "Security Models and Requirements for Healthcare Application Clouds". 3rd International Conference on Cloud Computing", IEEE, 2010.
[13] Coatrieux *et al*. "Relevance of Watermarking in Medical Imaging". Proceedings of the IEEE EMBS Conf. on Information Technology Applications in Biomedicine, Arlington, USA, pp 250-255, 2000.
[14] Aminzou *et al.* "Towards a Secure Access to Patient Data in Cloud Computing Environments". Security Days (JNS3), IEEE, 2013.
[15] Elgamal, Hikal & Abou-Chadi. "Secure Medical Images Sharing over Cloud Computing environment". International Journal of Advanced Computer Science and Applications (IJACSA), vol. 4, N°5, 2013.
[16] Bouslimi and Coatrieux. "A Joint/Encryption Watermarking System for Verifying the Reliability of Medical Images". Medical Data Privacy Handbook pp. 493-526, Springer, 2015.
[17] Al-Haj, Hussein and Abandah. "Combining Cryptography and Digital Watermarking for Secured Transmission of Medical Images". Second International Conference on Information Management (ICIM), IEEE; 2016.
[18] Garkotti *et al*. "Detection of Insider attacks in Cloud based e-healthcare". International Conference on Information Technology, IEEE, 2014.
[19] Lee *et a*l. "Implementation of MapReduce-based Image Conversion Module in Cloud Computing Environment". Proc. of Int. Conference on Advances in Computing, 2011.