Sistema de generación de Tokens para proveer seguridad a los datos sensibles de los clientes de establecimientos comerciales.

Trabajo Terminal No. 2020 - A112

Alumnos: López López Raquel Iris, Martínez Clavería Karen Olivia.*
Directores: Díaz Santiago Sandra, Rodríguez Henríquez Lil María Xibai
e-mail: sirioel16@gmail.com

Resumen – El presente trabajo terminal tiene como objetivo desarrollar un sistema que ofrezca el servicio de *tokenización* para proteger los datos sensibles intercambiados en una transacción comercial entre los compradores y el establecimiento comercial. Haciendo uso de este sistema los comercios o puntos de venta que reciben pagos con tarjetas podrán generar y administrar los tokens de los números de las tarjetas de los clientes que realicen pagos con este medio. El servicio pretende, por un lado, brindar privacidad al cliente, de modo que sus datos bancarios no queden expuestos y que ningún atacante externo o incluso entidades del punto de venta tengan acceso a estos, y así evitar el mal uso de ellos. Por otro lado, brinda a los comercios una alternativa de tokenización sin que ellos tengan que preocuparse por la implementación de seguridad.

Palabras clave – Ciencias de la computación, Criptografía, PCI DSS (Payment Card Industry Data Security Standard), Tokenización.

1. Introducción

Las transacciones comerciales son comúnmente llevadas a cabo mediante el uso de tarjetas bancarias por gran parte de la población mundial. Un banco, es un intermediario por el cual es posible intercambiar dinero desde una cuenta bancaria de un consumidor a alguna otra entidad, acarrea muchos beneficios a las partes involucradas, entre ellos la comodidad para realizar cobros y pagos, además de la reducción de tiempo. Otro de los beneficios presumibles por las entidades bancarias es el de proveer seguridad en el intercambio de dinero, ya que, al ofrecer medios de pago a través de productos como tarjetas de crédito, el cliente no requiere llevar con él dinero en efectivo. Sin embargo, esto no es del todo cierto, ya que la información de las tarjetas puede quedar expuesta en cada transacción, debido a que cada vez que se hace uso de ella en un establecimiento comercial, comparte datos sensibles o información valiosa para un posible atacante. Información como el número de tarjeta, la fecha de vencimiento y el CCV son requeridas para identificar y validar la transacción, por lo cual no pueden ser omitidos, sin embargo, algunas de las alternativas proponen sustituir estos valores por otros que los representen, sin que esto conlleve exponer información sensible.
Para tratar de solventar este problema, se han propuesto diversas soluciones, una de las más estudiadas ha sido la *tokenización*.
Un proceso de tokenización consiste en reemplazar información sensible con un valor llamado token, de tal forma que la información que se pretende resguardar no quede expuesta.
La tokenización puede implementarse a nivel de sofware, hardaware o servicio y además tiene la siguiente clasificación dada por la PCI SSC [1].

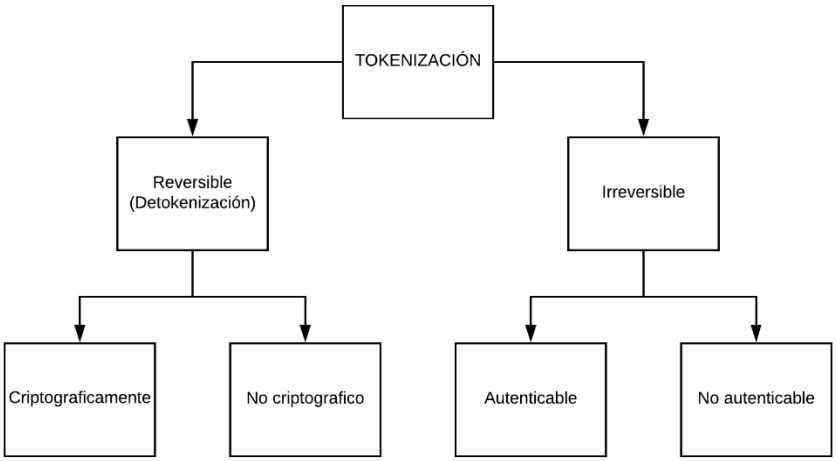

Figura 1. Clasificación de Tokenización

Reversible: en un proceso de tokenización reversible se puede ir en los dos sentidos, es decir, a partir de un dato se genera un token y luego, desde el token generado regresar al dato que fue tokenizado. Al proceso de regresar al dato original a través del token se le conoce como *detokenizacion*
La tokenización reversible puede construirse a través de técnicas criptográficas en los que se hace uso del cifrado, así como medios no criptográficos tales como la construcción de estructuras de datos donde se pueda relacionar el dato con el token asociado y el token con dato.

Irreversible: este proceso es unidireccional, ya que no se puede regresar al dato original a partir del token, ya que para ello se hace uso de alguna función matemática de un solo sentido (one-way function).
Si bien no se puede obtener el dato del que deriva el token, se puede acreditar que se utilizó un dato determinado para generarlo; si se puede saber esto, se dice que es un token irreversible autenticable.

La autenticación es uno de los servicios que ofrece la criptografía, la cual consiste en la verificación del origen o la prueba de la identidad de alguna entidad.
Se puede ver que, en ambas clasificaciones, (la tokenización reversible e irreversible) se hace uso de algún elemento de origen criptográfico, por lo que para poder llevar a cabo el desarrollo de un proceso de tokenización se hace uso de esta ciencia debido a que de hecho la tokenización es una aplicación criptográfica.

Sabiendo que un proceso de tokenización es una aplicación de la criptografía, se puede entender que los componentes de un sistema de tokenización involucran componentes de un esquema criptográfico.

Componentes de un sistema de Tokenización

1.- Método para generación de tokens: Algoritmo o procedimiento a seguir para obtener el token correspondiente a un dato dado.
2.- Procedimiento de mapeo de token.
3.- Repositorio que almacena pares de una cadena original y su respectivo Token.
4.- Administración de llave criptográfica (creación, uso, manejo y protección de llaves).
5.- Método de detokenización (en caso de token reversible).

Un proceso de tokenización debe cumplir los siguientes requerimientos:

1.- Preservación de formato: En determinados entornos, se requiere que cuando se cifra la información, esta conserve el mismo formato que los datos originales, el principal motivo es que al momento de almacenar la información en una base de datos no debe existir incompatibilidad de tipo o formato de datos [3].
2.- Unicidad: El método de generación de tokens deber ser determinista. Se refiere a que al momento de cifrar un dato en particular o en este caso al tratar de generar el token de un dato específico, se llegue a un único token y, además, cuando se generen tokens para dos datos diferentes el resultado deberá ser diferente para cada uno.

La tokenización se puede aplicar a varios campos, algunos ejemplos se listan a continuación.

- En transacciones comerciales, tokenizando información asociada a una tarjeta de crédito como el número de cuenta.
- En el sector de salud, tokenizando el número de seguridad social.
- En un proceso de votación, tokenizando los datos asociados a una identificación electoral.

Los ejemplos son muchos, pero lo que guardan en común es que en cualquiera de los casos la tokenización sirve para mantener la privacidad de datos sensibles, es decir, aquellos datos que requieren protección y para los cuales la ley establece un tratamiento especial, o información personal privada de un individuo, por ejemplo, ciertos datos personales y bancarios, contraseñas de correo electrónico e incluso el domicilio en algunos casos [4].

A continuación, se muestran algunos de los esquemas diseñados con anterioridad:

TÍTULO	CARACTERÍSTICAS	VENTAJAS	DESVENTAJAS
Generación de Tokens para proteger los datos de tarjetas bancarias (TT 2017-B008) [5]	♦ Servicio web de generación de Tokens ♦ Tienda que hace uso del servicio de Tokens	Preserva el formato, es decir, genera Tokens con el mismo alfabeto del número PAN asociado a una tarjeta.	Genera Tokens para una sola tienda.
A Cryptographic Study of Tokenization Systems (Paper) [6]	♦ Cifrado simétrico ♦ Cifrado con preservación de formato		
Solución propuesta	Servicio web que: Genera y administra Tokens para diferentes comercios	Genera Tokens para diversas tiendas.	

Tabla 1. Resumen de productos similares.

2. Objetivo

Objetivo general
Desarrollar un sistema web prototipo, que provea el servicio de tokenización a terceros para proteger datos sensibles de tarjetas bancarias.

Objetivos específicos:

- ♦ Analizar las diferentes propuestas de solución para la construcción de un producto de tokenización enfocado en el ambiente comercial para determinar el algoritmo o los algoritmos que preserven el formato, que puedan ser almacenados en un repositorio y cuya seguridad no haya sido vulnerada.
- ♦ Implementar un algoritmo de tokenización que provea privacidad en las transacciones comerciales entre los comerciantes y los compradores cuando estos hacen uso de su tarjeta bancaria en varios establecimientos.
- ♦ Desarrollar un sistema web que haga uso del algoritmo o los algoritmos analizados e implementados previamente desde donde los comercios puedan generar y administrar los tokens para el ocultamiento de datos de sus compradores.

3. Justificación

En la actualidad, los establecimientos comerciales son quienes se encargan de generar tokens para sus clientes de manera individual cada que se lleva a cabo el pago de algún servicio o bien. En este escenario, los que hacen la venta son los mismos que se encargan de la seguridad de los datos de las tarjetas bancarias de los compradores. Por ello, la problemática que pretende resolver este Trabajo Terminal consiste en ofrecer a los comercios y compradores una alternativa en la que la seguridad de los datos sensibles no quede bajo responsabilidad total del comercio y con ello, no puedan aprovecharse de esta ventaja y hacer mal uso de los datos. Además, de que, el que cuenten con un servicio externo que genere y administre los tokens les ahorraría costos, eficiencia, eficacia y la complejidad de almacenar y administrar todos los datos de sus clientes. El presente trabajo busca solventar dicha problemática desarrollando un sistema web que se encargue de realizar este trabajo para los establecimientos comerciales que lo requieran. Con el desarrollo de este sistema podemos proporcionar una mayor seguridad tanto para los clientes como para los corporativos, debido a que los empleados no tendrán acceso a los datos sensibles, así como la innovación porque se implementarán algoritmos recientes cuya seguridad está demostrada.
La aplicación de este trabajo se extiende a cualquier establecimiento comercial que acepte pagos por medio de tarjetas bancarias y que busque preservar la seguridad de sus clientes. Los beneficiarios serían tanto los comercios que no tendrían que preocuparse por los métodos de implementación para la generación de token y por otro lado los compradores, quienes podrán usar de manera segura sus tarjetas bancarias en dichos establecimientos.

4. Productos o Resultados esperados

Aplicación web (1) desde la cual se podrán generar y administrar tokens asociados a datos sensibles de compradores de distintos establecimientos comerciales. El servicio web hace uso del algoritmo generador de tokens (2), desde el cual se podrán producir los tokens asociados a cierta información sensible. Para relacionar y administrar los tokens y la información a la que están asociados, el generador de tokens hace uso de un repositorio (3) o base de datos relacional.
En los productos esperados se incluye la documentación técnica derivada del análisis, diseño y desarrollo de dicha aplicación web. Además del Manual de usuario del sistema.

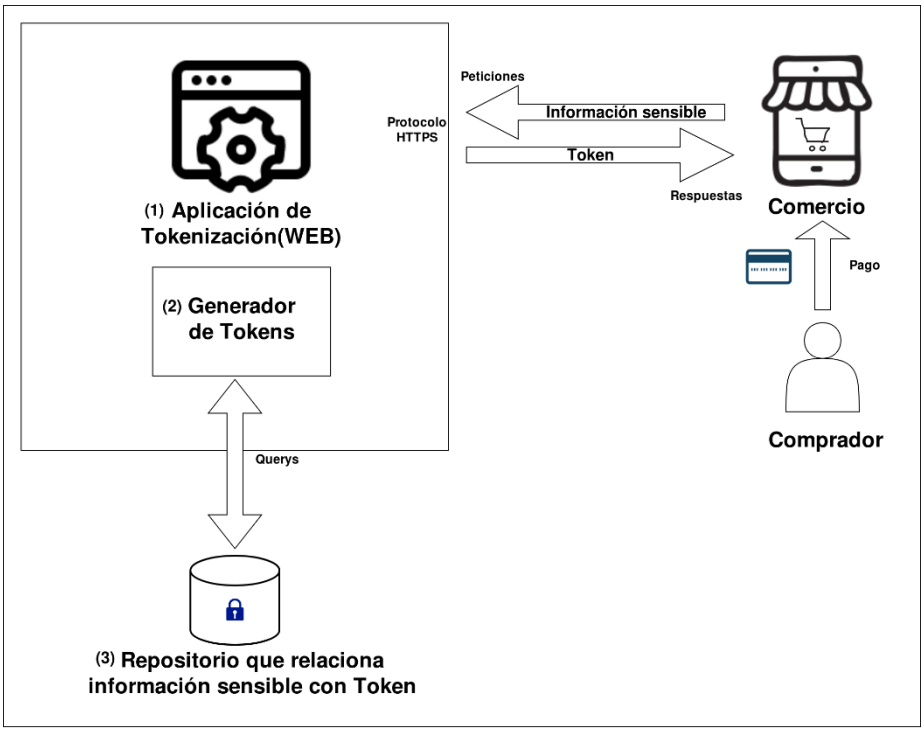

Figura 2. Arquitectura del sistema.

5. Metodología

Se empleará la metodología de software SCRUM [7], debido a que es incremental y permite que se puedan apreciar detalladamente las actividades, los tiempos destinados a cada módulo, los avances específicos en cada proceso iterativo y las funcionalidades del proyecto. Las etapas de cada Sprint son las mismas a diferencia del módulo al que se destina (Base de Datos, Página web, Método, etc.).

Fases de SCRUM	Descripción
1. Preparación del proyecto	Las estimaciones del Backlog. Se definirá un documento en el que se reflejarán los requisitos del sistema por prioridades. Se obtendrá además un Sprint Backlog, que es la lista de tareas
2. Planificación del Sprint	La Estimación del Sprint Reunión de Planificación (Sprint Panning Meeting)
3. Seguimiento del Sprint	Reuniones diarias para evaluar el avance (Sprint Daily Meeting). * ¿Qué trabajo se realizó desde la reunión anterior? * ¿Qué trabajo se hará hasta una nueva reunión? *Inconvenientes que han surgido y qué hay que solucionar para poder continuar.
4. Revisión del Sprint	Revisión del incremento, presentación de resultados finales. Reunión Revisión del Sprint (Sprint Review Meeting). Reunión de Retrospectiva (Sprint Retrospective Meeting).

Tabla 2: Etapas de cada Sprint

Cada fase de Sprint se llevará a cabo de acuerdo a la definición estipulada en la fase 1 y de acuerdo a la lista de tareas.

Pila del Sprint	Tareas	Tiempo	Estado
Sprint 0 Modelo de Negocio	Descripción del proceso de negocio		
	Definición de visión, objetivos y alcance		
Sprint 1 Análisis	Análisis de riesgo		
	Identificación de funcionalidades y restricciones		
	Requerimientos funcionales y no funcionales		
	Casos de uso		
Sprint 2 Diseño	Diseño de arquitectura		
	Diseño de componentes (BD, Algoritmo Generador, Aplicación, etc.)		
	Diseño de interfaces		
Sprint 3 Implementación	Desarrollo de módulo Generador de Tokens		
	Desarrollo de módulo de Base de Datos		
	Desarrollo de módulo Web		
	Integración del sistema		
Sprint 4 Pruebas	Ejecución de pruebas Unitarias		
	Ejecución de pruebas funcionales		
Sprint 5 Despliegue	Ejecución en el ambiente final		
Sprint 6 Documentación	Creación de manuales y reportes técnicos		

Tabla 3. Actividades del proyecto utilizando la metodología de software SCRUM.

En la planificación de cada Sprint, se estimará el tiempo de cada tarea y durante el seguimiento o ejecución del Sprint se actualizará el estado de cada tarea.

6. Cronograma

Anexo I.

7. Referencias

[1] Payment Card Industry Security Standards Council. Tokenization Product Security Guideless-Irreversible and Irreversible Tokens 2015. URL:https://www.pcisecuritystandards.org/documents/Tokenization_Product_Security_Guidelines.pdf.
[2] Alfred J. Menezes, Paul C. Van Oorschot, and Scott A. Vanstone. Handbook of Applied Cryptography. CRC Press, USA, 1996.
[3] V. Gayoso Martínez, L. Hernández Encinas, A. Martín Muñoz, J. M. de Fuentes, L. González Manzano, Cifrado de datos con preservación del formato. En: JNIC2015, 2015, ISBN 978-84-9773-742-5, págs. 110-115. URL: http://hdl.handle.net/10612/5679.
[4] Diario Oficial de la Federación, 2010, Ley general de protección de datos personales en posesión de sujetos obligados, art. 3ro fracc. IX y X, pág. 3. URL: http://www.diputados.gob.mx/LeyesBiblio/ref/lfpdppp.htm.
[5] Ayala Zamorano Daniel, Borbolla Palacio Laura Natalia & Quezada Figueroa Ricardo. (2017) Generación de Tokens para proteger los datos de tarjetas bancarias. (TT 2017-B008).
[6] Sandra Díaz-Santiago, Lil María RodríguezHenríquez y Debrup Chakraborty. <A Cryptographic Study of Tokenization Systems>. En: Int. J. Inf. Sec. 15.4 (2016), págs. 413-432. DOI: 10.1007/s10207-015-0313-x. URL: //DOI.ORG/10.1007/s10207-015-0313-x.
[7] Jeff Sutherland y Ken Schwaber, 2017, The SCRUM Guide, URL: https://www.scrumguides.org/

8. Alumnos y Directores

López López Raquel Iris. - Alumna de la carrera de Ing. En Sistemas Computacionales en la ESCOM, Especialidad Sistemas Boleta: 2011630461, Tel. 5510087204, email sirioel16@gmail.com

Firma:_____

Martínez Clavería Karen Olivia. - Alumna de la carrera de Ing. en Sistemas Computacionales en la ESCOM, Especialidad Sistemas, Boleta: 2015630272, Tel. 5561380498, email i.systems.c@gmail.com

Firma:_____

Sandra Díaz Santiago.- Doctorado en Ciencias en Computación (CINVESTAV-IPN, 2014). Maestría en Ciencias (Matemáticas) (UAM-Iztapalapa, 2005). Licenciatura en Computación (UAM-Iztapalapa, 1998). Profesor Titular en ESCOM (Departamento de Ciencias e Ingeniería de la Computación), desde 2004,
Áreas de Interés: Criptografía, Pseudoaleatoriedad, Seguridad Demostrable,
email: sdiazs@gmail.com,
sdiazsa@ipn.mx.

Firma:_____

Lil María Xibai Rodríguez Henríquez.-Doctorado en Ciencias en Computación (CINVESTAV-IPN, 2015).
Maestría en Ciencias de la Computación (CINVESTAV-IPN, 2009). Ingeniería en Computación (Universidad Albert Einstein, San Salvador, El Salvador, 2005).
Investigadora del CONACyT comisionada en el Instituto Nacional de Astrofísica Óptica y Electrónica (INAOE), desde 2016.
Áreas de Interés: Criptografía, búsqueda en texto cifrado, seguridad en bases de datos, seguridad demostrable.
Email: lmrodriguez@inaoep.mx

Firma:_____

Nombre del alumno(a): López López Raquel Iris TT No.: 2020 - A112

Título del TT: Sistema de generación de Tokens para proveer seguridad a los datos sensibles de los clientes de establecimientos comerciales.

ACTIVIDAD	AGO	SEP	OCT	NOV	DIC	ENE	FEB	MAR	ABR	MAY	JUN
Sprint 0 Modelo de Negocio											
Descripción del proceso de negocio											
Definición de visión, objetivos y alcance											
Sprint 1 Análisis											
Análisis de riesgo											
Identificación de funcionalidades y restricciones											
Requerimientos funcionales y no funcionales											
Casos de uso											
Sprint 2 Diseño											
Diseño de arquitectura											
Diseño de componentes (BD, Algoritmo Generador, Aplicación, etc.)											
Diseño de interfaces											
Evaluación de TT I											
Sprint 3 Implementación											
Desarrollo de módulo Generador de Tokens											
Desarrollo de módulo de Base de Datos											
Desarrollo de módulo Web											
Integración del sistema											
Sprint 4 Pruebas											
Ejecución de pruebas Unitarias											
Ejecución de pruebas funcionales											
Sprint 5 Despliegue											
Ejecución en el ambiente final											
Sprint 6 Documentación											
Creación de manuales y reportes técnicos											
Evaluación de TT II											

Nombre del alumno(a): Martínez Claveria Karen Olivia TT No.: 2020 - A112
Título del TT: Sistema de generación de Tokens para proveer seguridad a los datos sensibles de los clientes de establecimientos comerciales.

ACTIVIDAD	AGO	SEP	OCT	NOV	DIC	ENE	FEB	MAR	ABR	MAY	JUN
Sprint 0 Modelo de Negocio											
Descripción del proceso de negocio											
Definición de visión, objetivos y alcance											
Sprint 1 Análisis											
Análisis de riesgo											
Identificación de funcionalidades y restricciones											
Requerimientos funcionales y no funcionales											
Casos de uso											
Sprint 2 Diseño											
Diseño de arquitectura											
Diseño de componentes (BD, Algoritmo Generador, Aplicación, etc.)											
Diseño de interfaces											
Evaluación de TT I											
Sprint 3 Implementación											
Desarrollo de módulo Generador de Tokens											
Desarrollo de módulo de Base de Datos											
Desarrollo de módulo Web											
Integración del sistema											
Sprint 4 Pruebas											
Ejecución de pruebas Unitarias											
Ejecución de pruebas funcionales											
Sprint 5 Despliegue											
Ejecución en el ambiente final											
Sprint 6 Documentación											
Creación de manuales y reportes técnicos											
Evaluación de TT II											