

# GENERACIÓN DE TOKENS PARA PROTEGER LOS DATOS DE TARJETAS BANCARIAS

TRABAJO TERMINAL No. 2017-B008

PRESENTAN

DANIEL AYALA ZAMORANO

DAZ23AYALA@GMAIL.COM

LAURA NATALIA BORBOLLA PALACIOS

LN.BORBOLLA.42@GMAIL.COM

RICARDO QUEZADA FIGUEROA

QF7.RICARDO@GMAIL.COM

DIRECTORA

DRA. SANDRA DÍAZ SANTIAGO

CIUDAD DE MÉXICO, 9 DE MAYO DE 2018

ESCUELA SUPERIOR DE CÓMPUTO  
INSTITUTO POLITÉCNICO NACIONAL



# CONTENIDO

Planteamiento del problema

# CONTENIDO

Planteamiento del problema

Planteamiento de la solución

# CONTENIDO

Planteamiento del problema

Planteamiento de la solución

Algoritmos generadores de *tokens*

# CONTENIDO

Planteamiento del problema

Planteamiento de la solución

Algoritmos generadores de *tokens*

Conclusiones

# PLANTEAMIENTO DEL PROBLEMA, CONTENIDO

Planteamiento del problema

Planteamiento de la solución

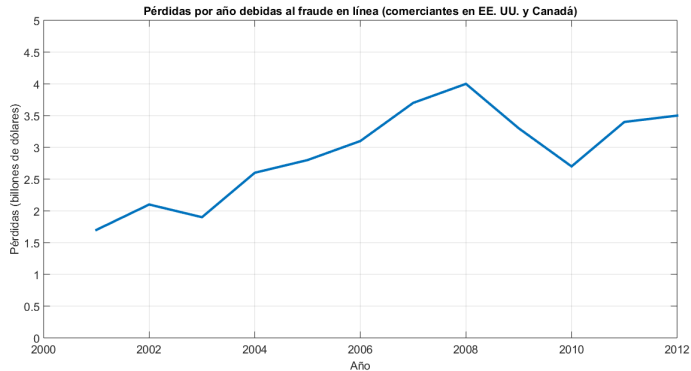
Algoritmos generadores de *tokens*

Conclusiones

# UN INICIO TORMENTOSO

- ▶ En la década de los 80 y 90, el comercio en línea comenzó a crecer y tomar importancia.
- ▶ Las empresas no estaban preparadas para el impacto que tuvieron y los fraudes relacionados con el comercio electrónico aumentaron rápidamente [1].
  - ▶ Visa y Mastercard reportaron, entre 1988 y 1998, pérdidas de 750 millones de dólares.

# UN INICIO TORMENTOSO



Pérdidas debidas al fraude en línea (2001-2012) [2].



# UN ESTÁNDAR PARA GOBERNARLOS A TODOS

- ▶ A inicios del 2000, las grandes compañías emisoras de tarjetas <sup>1</sup> comenzaron a publicar, individualmente, *buenas prácticas* de seguridad.
- ▶ Las empresas intentaron adoptar las prácticas, pero era tremendamente complicado y costoso.
- ▶ Se aliaron las compañías emisoras y, en 2004, publicaron un estándar unificado: PCI-DSS<sup>2</sup> [3].
  - ▶ Se hizo obligatorio para quienes realizasen más de 20K transacciones al año.
  - ▶ Tiene un gran número de requerimientos (y subrequerimientos), por lo que es difícil de satisfacer.

---

<sup>2</sup>VISA, MasterCard, American Express, entre otras.

<sup>2</sup>Payment Card Industry - Data Security Standard

# CAMBIO DE ESTRATEGIA

- ▶ Hasta ahora, el enfoque era proteger los datos sensibles donde sea que se encuentren y por donde sea que transiten.
- ▶ Surge un nuevo enfoque: cambiar la información valiosa, por *valores representativos* (tokens); es decir, la tokenización de la información.
- ▶ En 2011, el PCI-SSC <sup>3</sup> publicó las primeras guías para los procesos de tokenización [4].
  - ▶ Aunque indica lo que debe satisfacer el sistema tokenizador, no dice cómo generar los tokens.

---

<sup>3</sup>Payment Card Industry - Security Standards Council

# PERO ¿POR QUÉ?

A pesar de ser una práctica extendida, la tokenización sigue estando rodeada de desinformación y desconfianza.

- ▶ Se busca combatir la desinformación al estudiar e implementar cinco algoritmos tokenizadores, compararlos y mostrar los resultados.
- ▶ Hacer notar que la criptografía y la tokenización no están peleadas; pues la tokenización puede verse como una aplicación de la criptografía.

# PLANTEAMIENTO DE LA SOLUCIÓN, CONTENIDO

Planteamiento del problema

Planteamiento de la solución

Objetivos del proyecto	9
Metodología del proyecto	10
Prototipos	11

Algoritmos generadores de *tokens*

Conclusiones

# OBJETIVOS DEL PROYECTO




Lo que se busca con este proyecto es implementar un programa generador de *tokens* que provea confidencialidad a los datos de las tarjetas bancarias.

Además, con el afán de disminuir la desinformación existente sobre la tokenización, se busca obtener una comparativa de los algoritmos implementados.



# PROTOTIPOS

Este proyecto está dividido en 3 prototipos, los cuales son:

 <b>Prototipo de generación de tokens. ✓</b>	 <b>Prototipo de servicio web.</b>	 <b>Prototipo de tienda en línea.</b>
<p>Revisar e implementar diversos algoritmos generadores de tokens para hacer un programa tokenizador, así como realizar pruebas comparativas entre estos algoritmos.</p>	<p>Diseñar e implementar una API web capaz de comunicar al programa tokenizador con al menos una tienda en línea con el fin ofrecer el servicio de tokenización.</p>	<p>Implementar una tienda en línea que utilice la API web para poder revisar el correcto funcionamiento del servicio.</p>

## Prototipos del trabajo terminal.

# MARCO TEÓRICO, CONTENIDO

Planteamiento del problema

Planteamiento de la solución

Algoritmos generadores de *tokens*

Clasificación	13
---------------	----

Implementaciones	16
------------------	----

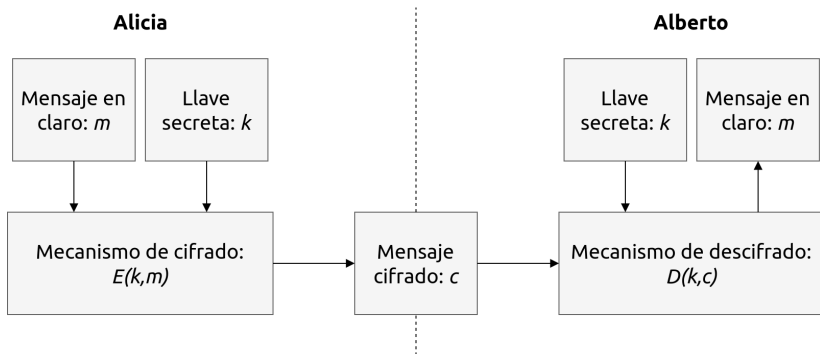
Resultados	18
------------	----

Conclusiones



# INTRODUCCIÓN A LA CRIPTOGRAFÍA

## CONFIDENCIALIDAD



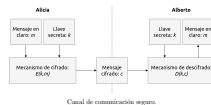
Canal de comunicación seguro.

# Generación de tokens para proteger los datos de tarjetas bancarias

## Marco teórico

### Introducción a la criptografía

### Introducción a la criptografía



La idea básica de la criptografía es transformar un mensaje para que solo las partes autorizadas puedan leerlo. En este caso Alicia es el emisor del mensaje y Alberto el receptor; ambos cuentan con un secreto común: la llave. Alicia usa la llave y el mecanismo de cifrado para transformar su mensaje en algo ilegible; Alberto utiliza la llave y el mecanismo de descifrado para obtener el mensaje original.

Esto se conoce como *confidencialidad*, y es el principal objetivo de la criptografía moderna. Los otros son *integridad*, *autenticación* y *no repudio*.

# GENERACIÓN DE NÚMEROS ALEATORIOS

Existen dos maneras de generar números aleatorios:

# GENERACIÓN DE NÚMEROS ALEATORIOS

Existen dos maneras de generar números aleatorios:

- ▶ Mediante un generador no determinístico  
(*Non-deterministic Random Bit Generator*, NRBG).

# GENERACIÓN DE NÚMEROS ALEATORIOS

Existen dos maneras de generar números aleatorios:

- ▶ Mediante un generador no determinístico  
(*Non-deterministic Random Bit Generator*, NRBG).
  - ▶ Están ligados a un proceso físico impredecible.

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Marco teórico

- └ Generación de números aleatorios

- └ Generación de números aleatorios

Existen dos maneras de generar números aleatorios:

- Mediante un generador no determinístico (*Non-deterministic Random Bit Generator*, NRBG).
- Están ligados a un proceso físico impredecible.

Como con los volados: si tuviéramos manera de medir con precisión todos los subprocesos involucrados (el tamaño y peso de la moneda, la forma de mis dedos, la posición de la moneda con respecto a mis dedos, el punto de impacto, la fuerza del impacto, la resistencia del aire, la fuerza de gravedad en el punto de la tierra en cuestión, la presión atmosférica, etc) los volados dejarían de ser aleatorios.

Estos generadores buscan este tipo de procesos para generar números. Por ejemplo, la temperatura del procesador en un momento dado, los tiempos de acceso a memoria, los tiempos de respuesta a través de una comunicación por internet, la posición del mouse, etcétera.

# GENERACIÓN DE NÚMEROS ALEATORIOS

Existen dos maneras de generar números aleatorios:

- ▶ Mediante un generador no determinístico  
(*Non-deterministic Random Bit Generator*, NRBG).
  - ▶ Están ligados a un proceso físico impredecible.
  - ▶ Son difíciles de implementar y no existe ningún estándar aprobado.

# Generación de tokens para proteger los datos de tarjetas bancarias

└ Marco teórico

└ Generación de números aleatorios

└ Generación de números aleatorios

Existen dos maneras de generar números aleatorios:

- Mediante un generador no determinístico (*Non-deterministic Random Bit Generator*, NRBG).
  - Están ligados a un proceso físico impredecible.
  - Son difíciles de implementar y no existe ningún estándar aprobado.

Dada su naturaleza dependiente del hardware, son difíciles de implementar y, por lo tanto, aún no existe ningún estándar al respecto.



# GENERACIÓN DE NÚMEROS ALEATORIOS

Existen dos maneras de generar números aleatorios:

- ▶ Mediante un generador no determinístico (*Non-deterministic Random Bit Generator*, NRBG).
  - ▶ Están ligados a un proceso físico impredecible.
  - ▶ Son difíciles de implementar y no existe ningún estándar aprobado.
- ▶ Mediante un generador determinístico (*Deterministic Random Bit Generator*, DRBG).

# GENERACIÓN DE NÚMEROS ALEATORIOS

Existen dos maneras de generar números aleatorios:

- ▶ Mediante un generador no determinístico (*Non-deterministic Random Bit Generator*, NRBG).
  - ▶ Están ligados a un proceso físico impredecible.
  - ▶ Son difíciles de implementar y no existe ningún estándar aprobado.
- ▶ Mediante un generador determinístico (*Deterministic Random Bit Generator*, DRBG).
  - ▶ Utiliza un mecanismo claramente definido para producir secuencias de bits a partir de un valor inicial.

# Generación de tokens para proteger los datos de tarjetas bancarias

└ Marco teórico

└ Generación de números aleatorios

└ Generación de números aleatorios

Existen dos maneras de generar números aleatorios:

- ▶ Mediante un generador no determinístico (*Non-deterministic Random Bit Generator, NRBG*).
  - ▶ Están ligados a un proceso físico impredecible.
  - ▶ Son difíciles de implementar y no existe ningún estándar aprobado.
- ▶ Mediante un generador determinístico (*Deterministic Random Bit Generator, DRBG*).
  - ▶ Utiliza un mecanismo claramente definido para producir secuencias de bits a partir de un valor inicial.

Este valor es conocido como semilla.

# GENERACIÓN DE NÚMEROS ALEATORIOS

Existen dos maneras de generar números aleatorios:

- ▶ Mediante un generador no determinístico (*Non-deterministic Random Bit Generator*, NRBG).
  - ▶ Están ligados a un proceso físico impredecible.
  - ▶ Son difíciles de implementar y no existe ningún estándar aprobado.
- ▶ Mediante un generador determinístico (*Deterministic Random Bit Generator*, DRBG).
  - ▶ Utiliza un mecanismo claramente definido para producir secuencias de bits a partir de un valor inicial.
  - ▶ Dada la naturaleza determinística, se dice que los números son *pseudoaleatorios*.

# Generación de tokens para proteger los datos de tarjetas bancarias

## └ Marco teórico

### └ Generación de números aleatorios

### └ Generación de números aleatorios

Existen dos maneras de generar números aleatorios:

- Mediante un generador no determinístico (*Non-deterministic Random Bit Generator*, NRBG).
  - Están ligados a un proceso físico impredecible.
  - Son difíciles de implementar y no existe ningún estándar aprobado.
- Mediante un generador determinístico (*Deterministic Random Bit Generator*, DRBG).
  - Utiliza un mecanismo claramente definido para producir secuencias de bits a partir de un valor inicial.
  - Dada la naturaleza determinística, se dice que los números son pseudoaleatorios.

Dada la naturaleza determinística, alguien con el mismo valor inicial (con la misma semilla) puede generar los mismos números pseudoaleatorios. En un criptográfico se busca evitar esto, por lo que el valor de la semilla se debe mantener en secreto.

Generalmente el valor inicial se obtiene a través de una fuente de aleatoriedad (o de entropía). Estas, al igual que los NRBG, están ligadas a un proceso físico, pero a diferencia de estos, con un valor pequeño pueden generar muchos números pseudoaleatorios.

# ALGORITMOS GENERADORES DE *tokens*, CONTENIDO

Planteamiento del problema

Planteamiento de la solución

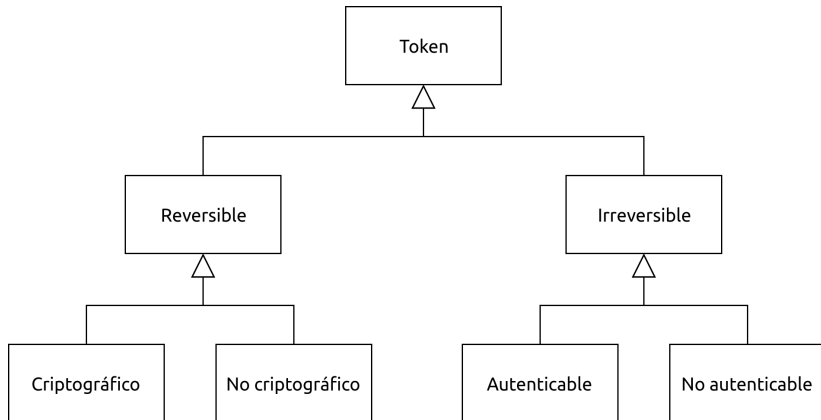
Algoritmos generadores de *tokens*

Conclusiones

Reporte de avances 21

Trabajo a futuro 22

# CLASIFICACIÓN DEL PCI SSC



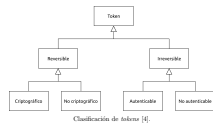
Clasificación de *tokens* [4].

## Generación de tokens para proteger los datos de tarjetas bancarias

- Algoritmos generadores de *tokens*

- Clasificación

- Clasificación del PCI SSC



Los irreversibles no pueden ser reconvertidos al PAN (de ninguna manera, mas que con fuerza bruta). Los autenticables funcionan como una función Hash: si tienes el PAN y el token, se puede validar que ese token es el par de ese PAN. Los no autenticables no pueden validar esto último.

Los reversibles permiten obtener el PAN a partir del token. Los no criptográficos ocupan funciones pseudoaleatorias y una base de datos para guardar las relaciones PAN-token. Los criptográficos ocupan un esquema de cifrado tradicional: un PAN mas una llave permiten obtener un token; la llave y el token pueden ser ocupados para obtener el PAN. No se ocupa una base de datos.



# CLASIFICACIÓN DEL PCI SSC

¿«*No criptográficos*»?

La clasificación anterior presenta los siguientes problemas:

- ▶ A pesar del nombre, los *no criptográficos* ocupan diversas aplicaciones de la criptografía para generar *tokens*.

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
- └ Clasificación
- └ Clasificación del PCI SSC

La clasificación anterior presenta los siguientes problemas:

- A pesar del nombre, los no criptográficos ocupan diversas aplicaciones de la criptografía para generar tokens.

Por ejemplo, la justificación para los no autenticables es para dar soporte a aplicaciones obsoletas que necesitan un formato de PAN válido. Esto se puede lograr con los no criptográficos sin guardar nada en la base; o pasando puros ceros en el campo del PAN.

El caso para los autenticables permite verificar la tarjeta del cliente en una compra cuando este perdió el comprobante. En est caso no resulta claro por qué la tienda (o el sistema tokenizador) no guardaría la transacción original.

# CLASIFICACIÓN DEL PCI SSC

¿«*No criptográficos*»?

La clasificación anterior presenta los siguientes problemas:

- ▶ A pesar del nombre, los *no criptográficos* ocupan diversas aplicaciones de la criptografía para generar *tokens*.
- ▶ Los casos de uso que el PCI SSC prevé en [4] para los irreversibles resultan artificiosos.

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*

- └ Clasificación

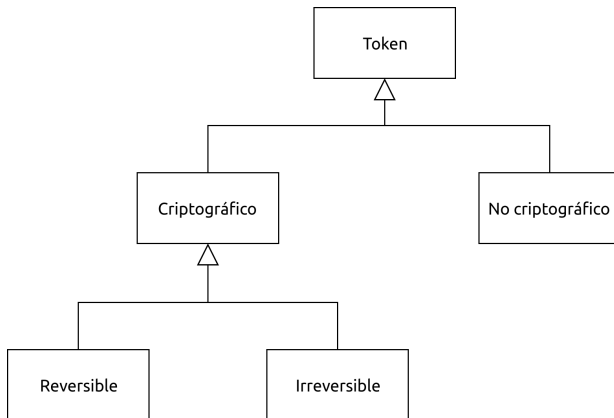
- └ Clasificación del PCI SSC

La clasificación anterior presenta los siguientes problemas:

- A pesar del nombre, los no criptográficos ocupan diversas aplicaciones de la criptografía para generar *tokens*.
- Los casos de uso que el PCI SSC prevé en [4] para los *token* resultan artificiales.

El problema con el PCI es que parecen pensar que la criptografía se limita a esquemas tradicionales, en donde hay una llave. La generación de números pseudoaleatorios seguros es también una aplicación de la criptografía.

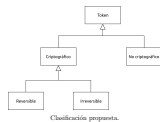
# CLASIFICACIÓN PROPUESTA



Clasificación propuesta.

## Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
  - └ Clasificación
    - └ Clasificación propuesta



Los únicos que se contemplan como «no criptográficos» son los que están basados en generadores realmente aleatorios. Todos los demás caen en la categoría de «criptográficos». Los reversibles son los que están basados en esquemas tradicionales (v. gr. los cifrados que preservan el formato). Los irreversibles necesitan de una base de datos para poder hacer el proceso inverso.

# ALGORITMOS IMPLEMENTADOS

- Reversibles:

# ALGORITMOS IMPLEMENTADOS

- ▶ Reversibles:
  - ▶ FFX (*Format-preserving Feistel-based Encryption*).  
Publicado por Mihir Bellare, Phillip Rogaway y Terence Spies en [6].



# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
  - └ Implementaciones
    - └ Algoritmos implementados

- Reversibles:
  - FFX (Format-preserving Feistel-based Encryption).  
Publicado por Mikar Bodden, Phillip Rogaway y Tsvi  
Spies en [6].

Es una propuesta de estándar para el NIST. Los autores son los principales precursores de los cifrados que preservan el formato.

El método está basado en redes Feistel y una función de ronda que ocupa CBC-MAC-AES.

# ALGORITMOS IMPLEMENTADOS

- ▶ Reversibles:
  - ▶ FFX (*Format-preserving Feistel-based Encryption*).  
Publicado por Mihir Bellare, Phillip Rogaway y Terence Spies en [6].
  - ▶ BPS. Publicado por Eric Brier, Thomas Peyrin y Jacques Stern en [7].

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
  - └ Implementaciones
    - └ Algoritmos implementados

- Reversibles:
  - FFX (*Format-preserving Feistel-based Encryption*).  
Publicado por Mihir Bellare, Phillip Rogaway y Terence Spies en [6].
  - BPS. Publicado por Eric Brier, Thomas Peyrin y Jacques Stern en [7].

También es propuesta de estándar para el NIST. Representa la principal competencia de FFX.

Al igual que FFX, ocupa redes Feistel de forma interna; se diferencian en algunos detalles de instanciación y en que BPS está diseñado para cadenas de longitud arbitraria.

# ALGORITMOS IMPLEMENTADOS

- ▶ Reversibles:
  - ▶ FFX (*Format-preserving Feistel-based Encryption*).  
Publicado por Mihir Bellare, Phillip Rogaway y Terence Spies en [6].
  - ▶ BPS. Publicado por Eric Brier, Thomas Peyrin y Jacques Stern en [7].
- ▶ Irreversibles:

# ALGORITMOS IMPLEMENTADOS

- ▶ Reversibles:
  - ▶ FFX (*Format-preserving Feistel-based Encryption*).  
Publicado por Mihir Bellare, Phillip Rogaway y Terence Spies en [6].
  - ▶ BPS. Publicado por Eric Brier, Thomas Peyrin y Jacques Stern en [7].
- ▶ Irreversibles:
  - ▶ TKR. Publicado por Sandra Díaz-Santiago, Lil María Rodríguez-Henríquez y Debrup Chakraborty en [8].

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
  - └ Implementaciones
    - └ Algoritmos implementados

- Reversibles:
  - FFX (*Format-preserving Feistel-based Encryption*). Publicado por Mihir Bellare, Phillip Rogaway y Teroese Spies en [6].
  - RPS. Publicado por Eric Brier, Thomas Peyrin y Jacques Stern en [7].
- Irreversibles:
  - TKR. Publicado por Sandra Díaz-Santiago, Lil María Rodríguez-Henríquez y Delrap Chakraborty en [8].

El documento es el primer análisis formal sobre la generación de tokens. TKR es el primer método propuesto (cuya seguridad está formalmente probada) que no es un cifrado que preserva el formato.

# ALGORITMOS IMPLEMENTADOS

- ▶ Reversibles:
  - ▶ FFX (*Format-preserving Feistel-based Encryption*). Publicado por Mihir Bellare, Phillip Rogaway y Terence Spies en [6].
  - ▶ BPS. Publicado por Eric Brier, Thomas Peyrin y Jacques Stern en [7].
- ▶ Irreversibles:
  - ▶ TKR. Publicado por Sandra Díaz-Santiago, Lil María Rodríguez-Henríquez y Debrup Chakraborty en [8].
  - ▶ AHR (Algoritmo Híbrido Reversible). Longo, Aragona y Sala en [9].

# ALGORITMOS IMPLEMENTADOS

- ▶ Reversibles:
  - ▶ FFX (*Format-preserving Feistel-based Encryption*). Publicado por Mihir Bellare, Phillip Rogaway y Terence Spies en [6].
  - ▶ BPS. Publicado por Eric Brier, Thomas Peyrin y Jacques Stern en [7].
- ▶ Irreversibles:
  - ▶ TKR. Publicado por Sandra Díaz-Santiago, Lil María Rodríguez-Henríquez y Debrup Chakraborty en [8].
  - ▶ AHR (Algoritmo Híbrido Reversible). Longo, Aragona y Sala en [9].
  - ▶ DRBG (*Deterministic Random Bit Generator*). Adaptación a partir del estándar del NIST (*National Institute of Standards and Technology*) [10].



# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
  - └ Implementaciones
    - └ Algoritmos implementados

- ▶ Reversibles:
  - ▶ FFX (*Format-preserving Feistel-based Encryption*). Publicado por Mihir Bellare, Phillip Rogaway y Teroese Spies en [6].
  - ▶ RPS. Publicado por Eric Brier, Thomas Peyrin y Jacques Stern en [7].
- ▶ Irreversibles:
  - ▶ TKR. Publicado por Sandra Díaz-Santiago, Lil María Rodríguez-Henríquez y Delarup Chakraborty en [8].
  - ▶ AHR (*Algoritmo Híbrido Reversible*). Longo, Aragona y Sala en [9].
  - ▶ DRBG (*Deterministic Random Bit Generator*). Adaptación a partir del estándar del NIST (*National Institute of Standards and Technology*) [10].

En la gran mayoría de los casos se buscó no hacer implementaciones propias de primitivas criptográficas, sin embargo, en el caso del generador, se hizo un excepción, para darle un poco más de contenido al trabajo. Esto último dado que hacer un generador implica también validarlo con pruebas de aleatoriedad del NIST.

# DISEÑO DE PROGRAMA

## COMPONENTES

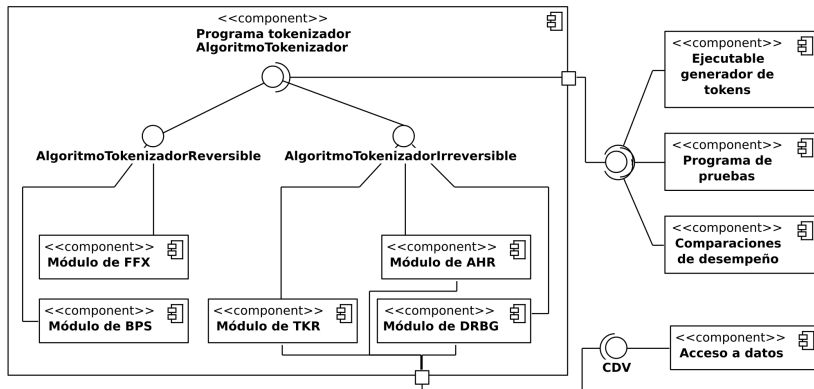


Diagrama de componentes del programa.

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
  - └ Implementaciones
    - └ Diseño de programa

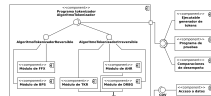


Diagrama de componentes del programa.

Se muestra la estructura interna del componente del programa tokenizador. Aunque adentro de este hay varios módulos, a las entidades externas solo les interesa comunicarse a través de la interfaz. El acceso a datos se maneja también a través de una interfaz externa al componente; los métodos irreversibles deben tener acceso a esta.

# RESULTADOS

## COMPARACIONES DE DESEMPEÑO

Las pruebas mostradas a continuación se realizaron en una computadora Toshiba S55-B5289 con las siguientes especificaciones:

- ▶ Procesador Intel Core i7-4710HQ.
  - ▶ 6M caché, hasta 3.50GHz.
  - ▶ 8 núcleos.
- ▶ 8GB de RAM.
- ▶ En los casos pertinentes, se utilizó AES-NI<sup>4</sup>.
- ▶ Se utilizó el compilador GCC versión 7.3.1.
- ▶ Se utilizó GNU gprof 2.30 como herramienta para obtener los perfiles de las funciones.

---

<sup>4</sup>*Intel Advanced Encryption Standard New Instructions.*

# Generación de tokens para proteger los datos de tarjetas bancarias

└ Algoritmos generadores de *tokens*

└ Resultados

└ Resultados

## RESULTADOS

### COMPARACIONES DE DESDESPROTO

Las pruebas mostradas a continuación se realizaron en una computadora Toshiba S55-B5289 con las siguientes especificaciones:

- Procesador Intel Core i7-4710HQ.
  - 4M caché, hasta 3.50GHz.
  - 4 núcleos.
- 8GB de RAM.
- En los casos pertinentes, se utilizó AES-NI<sup>4</sup>.
- Se utilizó el compilador GCC versión 7.3.1.
- Se utilizó GNU gprof 2.30 como herramienta para obtener los perfiles de las funciones.

---

<sup>4</sup>Intel Advanced Encryption Standard New Instructions.

Los tiempos de los reversibles son mucho más cortos.

Las gráficas muestran solo los procesos de tokenización: con la detokenización pasan cosas bastante similares.

# RESULTADOS

## COMPARACIONES DE DESEMPEÑO

	Tokenización	Detokenización
BPS	0.11 <i>ms</i>	0.11 <i>ms</i>
FFX	0.12 <i>ms</i>	0.12 <i>ms</i>
TKR	0.03 <i>ms</i>	0.00 <i>ms</i> <sup>4</sup>
AHR	48.16 <i>ms</i>	5.80 <i>ms</i>
DRBG	118.24 <i>ms</i>	10.84 <i>ms</i>

Comparación de tiempos de tokenización.

---

<sup>4</sup>Debido a la resolución de GNU-profiler, la operación de tokenización de TKR parece ser, incorrectamente, instánea.

# Generación de tokens para proteger los datos de tarjetas bancarias

└ Algoritmos generadores de *tokens*

└ Resultados

└ Resultados

Los tiempos de los reversibles son mucho más cortos.

Las gráficas muestran solo los procesos de tokenización: con la detokenización pasan cosas bastante similares.

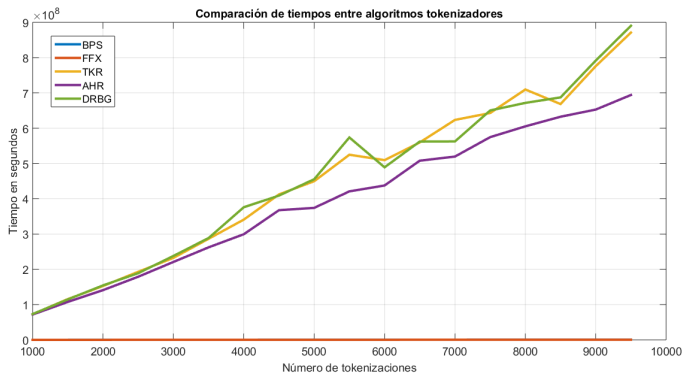
	Tokenización	Detokenización
BPS	0.11 ms	0.11 ms
PEX	0.12 ms	0.12 ms
TEB	0.02 ms	0.00 ms*
AHR	48.16 ms	3.50 ms
DRBG	118.24 ms	10.84 ms

Comparación de tiempos de tokenización.

\*Debido a la resolución de GNU-profiler, la operación de tokenización de TEB puede ser, incorrectamente, nula.

# RESULTADOS

## COMPARACIONES DE DESEMPEÑO



Tiempos de tokenización generales.



2018-05-02

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*

- └ Resultados

- └ Resultados

Los tiempos de los reversibles son mucho más cortos.

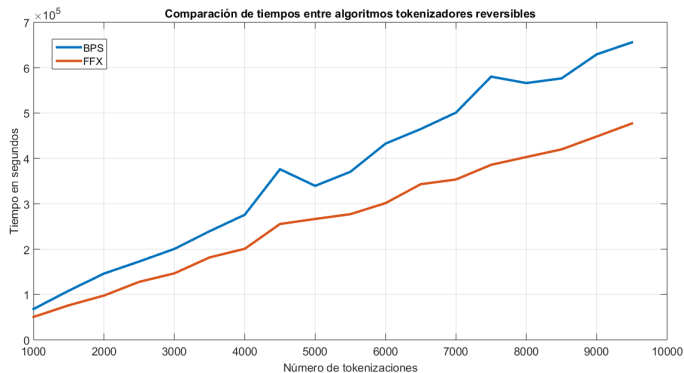
Las gráficas muestran solo los procesos de tokenización: con la detokenización pasan cosas bastante similares.

RESULTADOS  
COMPARACIONES DE DESERCIÓN



# RESULTADOS

## COMPARACIONES DE DESEMPEÑO



Tiempos de tokenización de reversibles.

2018-05-02

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
- └ Resultados
- └ Resultados

RESULTADOS  
Comparaciones de desempeño

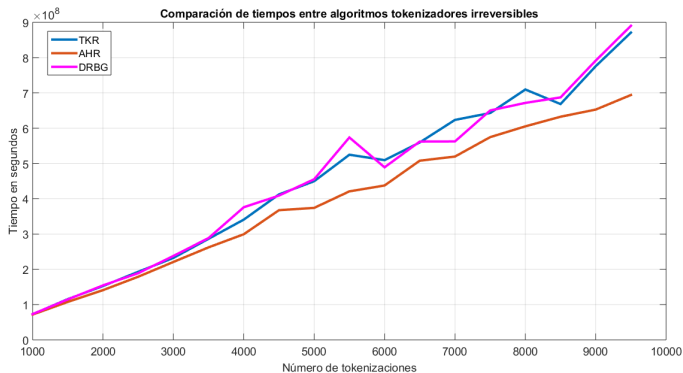


Los tiempos de los reversibles son mucho más cortos.

Las gráficas muestran solo los procesos de tokenización: con la detokenización pasan cosas bastante similares.

# RESULTADOS

## COMPARACIONES DE DESEMPEÑO



Tiempos de tokenización de irreversibles.

2018-05-02

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*

- └ Resultados

- └ Resultados

Los tiempos de los reversibles son mucho más cortos.

Las gráficas muestran solo los procesos de tokenización: con la detokenización pasan cosas bastante similares.

RESULTADOS  
Comparaciones de desempeño



Tiempo de tokenización de irreversibles.

# RESULTADOS

## PRUEBAS DE ALEATORIEDAD

En [11] el NIST describe un conjunto de pruebas estadísticas que sirven para determinar la aleatoriedad de un generador pseudoaleatorio. Se trata de 15 pruebas generales (algunas de ellas con subpruebas) que es necesario ejecutar sobre los bits generados.

Para cada instancia de los generadores implementados se ejecutó el conjunto de pruebas 20 veces, cada una con medio millón de bits (un total de veinte millones).

Resultados para generador basado en función hash:

- ▶ 112 bits de seguridad: 14 de 15.
- ▶ 128 bits de seguridad: 15 de 15.
- ▶ 192 bits de seguridad: 14 de 15.
- ▶ 256 bits de seguridad: 15 de 15.

# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
  - └ Resultados
    - └ Resultados

## RESULTADOS

### PRUEBAS DE ALEATORIEDAD

En [11] el NIST describe un conjunto de pruebas estadísticas que sirven para determinar la aleatoriedad de un generador pseudorrandatorio. Se trata de 15 pruebas generales (algunas de ellas con subpruebas) que es necesario ejecutar sobre los bits generados.

Para cada instancia de los generadores implementados se ejecutó el conjunto de pruebas 20 veces, cada una con medio millón de bits (un total de veinte millones).

Resultados para generador basado en función hash:

- ▶ 112 bits de seguridad: 14 de 15.
- ▶ 128 bits de seguridad: 15 de 15.
- ▶ 192 bits de seguridad: 14 de 15.
- ▶ 256 bits de seguridad: 15 de 15.

Estrictamente hablando, el generador basado en una función hash no es totalmente aleatorio, dado que falló en un par de pruebas. Sin embargo, el número de veces que se ejecutó el conjunto de pruebas (veinte) es un número relativamente pequeño (en comparación con lo recomendado por el NIST); esto por los recursos de cómputo que las pruebas exigen.

# RESULTADOS

## PRUEBAS DE ALEATORIEDAD

En [11] el NIST describe un conjunto de pruebas estadísticas que sirven para determinar la aleatoriedad de un generador pseudoaleatorio. Se trata de 15 pruebas generales (algunas de ellas con subpruebas) que es necesario ejecutar sobre los bits generados.

Para cada instancia de los generadores implementados se ejecutó el conjunto de pruebas 20 veces, cada una con medio millón de bits (un total de veinte millones).

Resultados para generador basado en cifrador por bloques:

- ▶ 112 bits de seguridad: 15 de 15.
- ▶ 128 bits de seguridad: 15 de 15.
- ▶ 192 bits de seguridad: 15 de 15.
- ▶ 256 bits de seguridad: 15 de 15.



# Generación de tokens para proteger los datos de tarjetas bancarias

- └ Algoritmos generadores de *tokens*
  - └ Resultados
    - └ Resultados

## RESULTADOS

### PRUEBAS DE ALEATORIEDAD

En [11] el NIST describe un conjunto de pruebas estadísticas que sirven para determinar la aleatoriedad de un generador pseudorandom. Se trata de 15 pruebas generales (algunas de ellas con subpruebas) que es necesario ejecutar sobre los bits generados.

Para cada instancia de los generadores implementados se ejecutó el conjunto de pruebas 20 veces, cada una con medio millón de bits (un total de veinte millones).

Resultados para generador basado en cifrador por bloques:

- ▶ 112 bits de seguridad: 15 de 15.
- ▶ 128 bits de seguridad: 15 de 15.
- ▶ 192 bits de seguridad: 15 de 15.
- ▶ 256 bits de seguridad: 15 de 15.

Estrictamente hablando, el generador basado en una función hash no es totalmente aleatorio, dado que falló en un par de pruebas. Sin embargo, el número de veces que se ejecutó el conjunto de pruebas (veinte) es un número relativamente pequeño (en comparación con lo recomendado por el NIST); esto por los recursos de cómputo que las pruebas exigen.

# CONCLUSIONES, CONTENIDO

Planteamiento del problema

Planteamiento de la solución

Algoritmos generadores de *tokens*

Conclusiones

# REPORTE DE AVANCES

# REPORTE DE AVANCES

- ▶ Primer prototipo: programa generador de *tokens* para dar confidencialidad a los datos de tarjetas bancarias.

# REPORTE DE AVANCES

- ▶ Primer prototipo: programa generador de *tokens* para dar confidencialidad a los datos de tarjetas bancarias.
  - ▶ Estudio de aspectos de la criptografía relacionados.

# REPORTE DE AVANCES

- ▶ Primer prototipo: programa generador de *tokens* para dar confidencialidad a los datos de tarjetas bancarias.
  - ▶ Estudio de aspectos de la criptografía relacionados.
  - ▶ Estudio de estándares y recomendaciones asociadas al tema.

# REPORTE DE AVANCES

- ▶ Primer prototipo: programa generador de *tokens* para dar confidencialidad a los datos de tarjetas bancarias.
  - ▶ Estudio de aspectos de la criptografía relacionados.
  - ▶ Estudio de estándares y recomendaciones asociadas al tema.
  - ▶ Análisis, diseño e implementación de algoritmos tokenizadores.

# REPORTE DE AVANCES

- ▶ Primer prototipo: programa generador de *tokens* para dar confidencialidad a los datos de tarjetas bancarias.
  - ▶ Estudio de aspectos de la criptografía relacionados.
  - ▶ Estudio de estándares y recomendaciones asociadas al tema.
  - ▶ Análisis, diseño e implementación de algoritmos tokenizadores.
- ▶ Comparación de desempeño entre algoritmos.



# REPORTE DE AVANCES

- ▶ Primer prototipo: programa generador de *tokens* para dar confidencialidad a los datos de tarjetas bancarias.
  - ▶ Estudio de aspectos de la criptografía relacionados.
  - ▶ Estudio de estándares y recomendaciones asociadas al tema.
  - ▶ Análisis, diseño e implementación de algoritmos tokenizadores.
- ▶ Comparación de desempeño entre algoritmos.
- ▶ Generador de números pseudoaleatorios junto con pruebas estadísticas de aleatoriedad.

# TRABAJO A FUTURO

## TRABAJO TERMINAL II

# TRABAJO A FUTURO

## TRABAJO TERMINAL II

- Prototipo dos: interfaz en red que permita comunicarse con el programa tokenizador.

# TRABAJO A FUTURO

## TRABAJO TERMINAL II

- ▶ Prototipo dos: interfaz en red que permita comunicarse con el programa tokenizador.
- ▶ Prototipo tres: tienda en línea que use de la interfaz en red.

# BIBLIOGRAFÍA I

- [1] SearchSecurity Staff. *The history of the PCI DSS standard: A visual timeline*. 2013. URL: <https://searchsecurity.techtarget.com/feature/The-history-of-the-PCI-DSS-standard-A-visual-timeline> (vid. pág. 7).
- [2] John S. Kiernan. *Credit Card And Debit Card Fraud Statistics*. 2017. URL: <https://wallethub.com/edu/credit-debit-card-fraud-statistics/25725/> (vid. pág. 8).
- [3] Payment Card Industry Security Standards Council. *Data Security Standard - Version 3.2*. 2016. URL: [https://www.pcisecuritystandards.org/documents/pci\\_dss\\_v3-2.pdf](https://www.pcisecuritystandards.org/documents/pci_dss_v3-2.pdf) (vid. pág. 9).

## BIBLIOGRAFÍA II

- [4] Payment Card Industry Security Standards Council. *Tokenization Product Security Guidelines – Irreversible and Reversible Tokens*. 2015. URL: [https://www.pcisecuritystandards.org/documents/Tokenization\\_Product\\_Security\\_Guidelines.pdf](https://www.pcisecuritystandards.org/documents/Tokenization_Product_Security_Guidelines.pdf) (vid. págs. 10, 17, 19, 21).
- [5] Microsoft. *Security Development Lifecycle*. 2008. URL: <https://www.microsoft.com/en-us/sdl/default.aspx> (vid. pág. 14).
- [6] Mihir Bellare, Phillip Rogaway y Terence Spies. “The FFX Mode of Operation for Format-Preserving Encryption”. Ver. 1.0. En: (2009) (vid. págs. 25, 26, 28, 30, 31, 33, 34).
- [7] Eric Brier, Thomas Peyrin y Jacques Stern. “BPS: a Format-Preserving Encryption Proposal”. En: (2010) (vid. págs. 25, 26, 28, 30, 31, 33, 34).

# BIBLIOGRAFÍA III

- [8] Sandra Diaz-Santiago, Lil María Rodríguez-Henríquez 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: <https://doi.org/10.1007/s10207-015-0313-x> (vid. págs. 25, 26, 28, 30, 31, 33, 34).
- [9] Riccardo Aragona, Riccardo Longo y Massimiliano Sala. “Several proofs of security for a tokenization algorithm”. En: *Appl. Algebra Eng. Commun. Comput.* 28.5 (2017), págs. 425-436. DOI: 10.1007/s00200-017-0313-3. URL: <https://doi.org/10.1007/s00200-017-0313-3> (vid. págs. 25, 26, 28, 30, 31, 33, 34).

## BIBLIOGRAFÍA IV

- [10] Elaine Barker y John Kelsey. *NIST Special Publication 800-90A - Recommendation for Random Number Generation Using Deterministic Random Bit Generators*. 2015. URL:  
<http://dx.doi.org/10.6028/NIST.SP.800-90Ar1>  
(vid. págs. 25, 26, 28, 30, 31, 33, 34).
- [11] Andrew Rukhin, Juan Soto, James Nechvatal y col. *NIST Special Publication 800-22 Revision 1a - A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*. 2010. URL:  
<https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf> (vid. págs. 48, 50).



# GENERACIÓN DE TOKENS PARA PROTEGER LOS DATOS DE TARJETAS BANCARIAS

TRABAJO TERMINAL No. 2017-B008

PRESENTAN

DANIEL AYALA ZAMORANO

DAZ23AYALA@GMAIL.COM

LAURA NATALIA BORBOLLA PALACIOS

LN.BORBOLLA.42@GMAIL.COM

RICARDO QUEZADA FIGUEROA

QF7.RICARDO@GMAIL.COM

DIRECTORA

DRA. SANDRA DÍAZ SANTIAGO

CIUDAD DE MÉXICO, 9 DE MAYO DE 2018

ESCUELA SUPERIOR DE CÓMPUTO  
INSTITUTO POLITÉCNICO NACIONAL

