Un vistazo a la tokenización

Presentan

Daniel Ayala Zamorano

daz23ayala@gmail.com

Laura Natalia Borbolla Palacios

ln.borbolla.42@gmail.com

Ricardo Quezada Figueroa

qf7.ricardo@gmail.com

Sandra Díaz Santiago

sdiazs@gmail.com

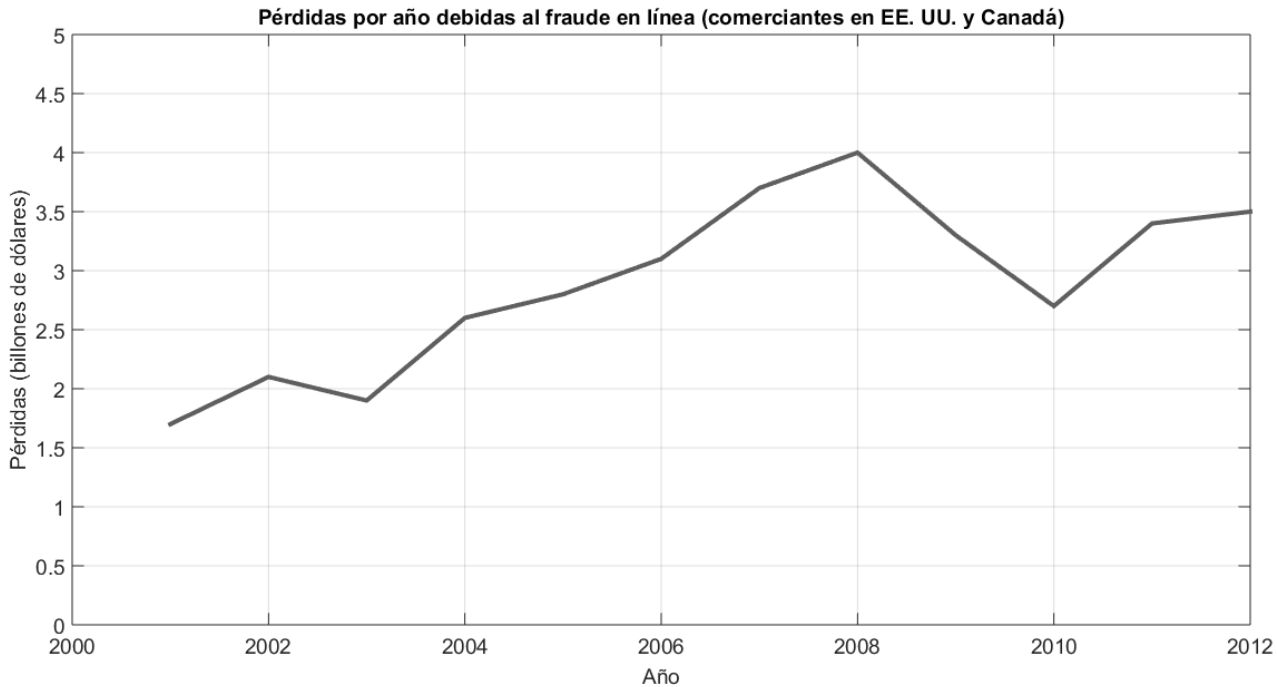
Primera Reunión de Ciberseguridad para la Industria 4.0
Puebla, 14 de octubre de 2018

Escuela Superior de Cómputo
Instituto Politécnico Nacional

Contenido

El problema de la protección de datos bancarios

- El crecimiento del comercio en línea, aunado a sistemas débilmente protegidos, propició un incremento en los robos de datos bancarios.



Pérdidas debidas al fraude en línea (2001-2012) [1].

El problema de la protección de datos bancarios

- En el 2004 se publicó el PCI DSS[1] [2].
- Hasta este momento el enfoque era proteger la información en donde sea que se encuentre.
- A pesar de la publicación del estándar, las filtraciones de datos no han cesado.

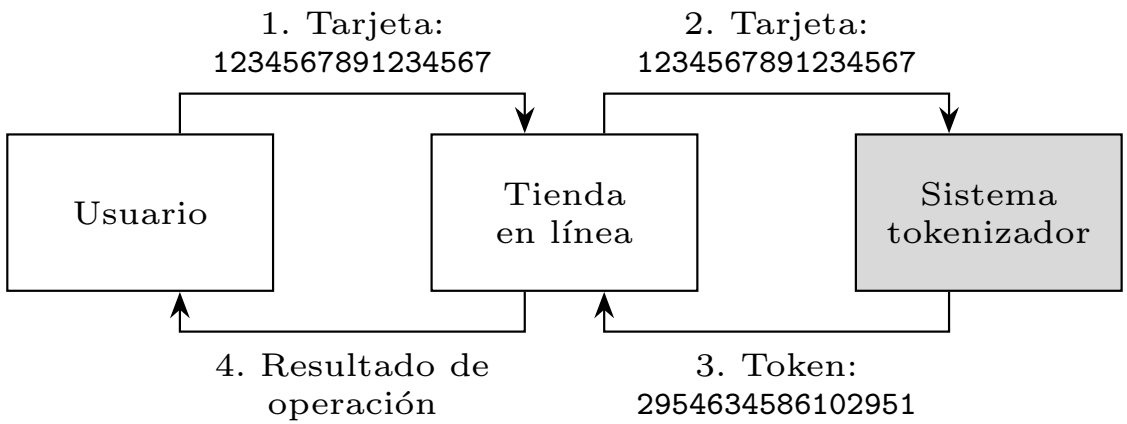
[1] *Payment Card Industry, Data Security Standard*

La tokenización en otros contextos

- Moneda de uso particular sin valor legal.
- Componente de seguridad en la comunicación por sesiones.
- Componente léxico de una gramática.
- Una unidad lingüística básica.
- Fenómeno social.

¿Qué es la tokenización?

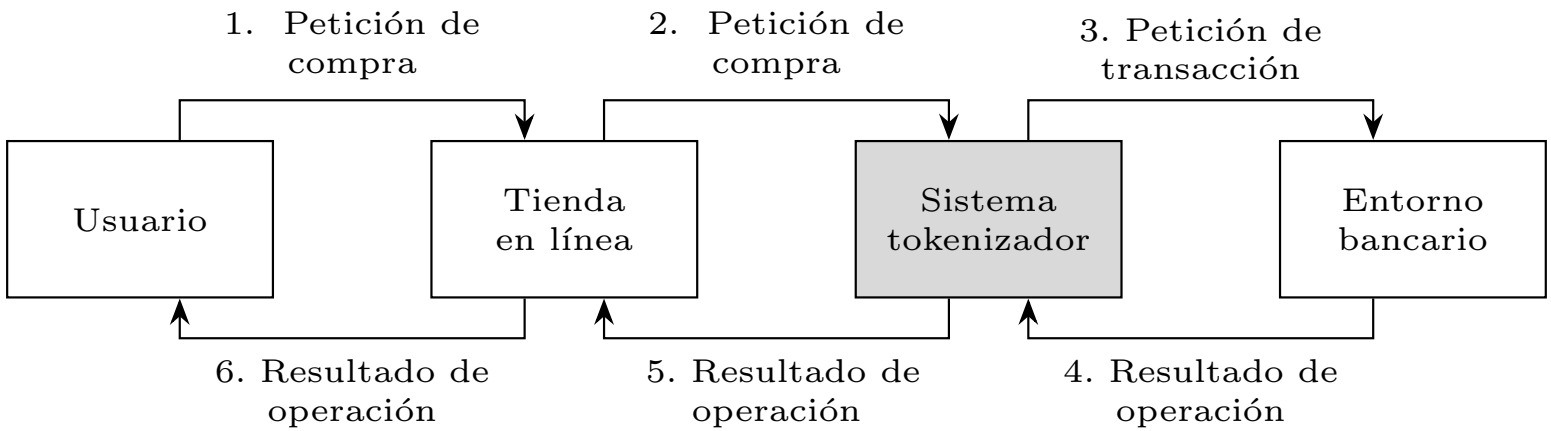
- Es la sustitución de datos sensibles por valores representativos sin una relación directa.
- Existen muchas empresas que proveen el servicio de tokenización, pero lo hacen sin detallar la forma en la que se realiza [3]-[5].
- En 2011, el PCI publicó su guía de tokenización [6].



Arquitectura de sistema tokenizador: operación de tokenización.

¿Qué es la tokenización?

- Es la sustitución de datos sensibles por valores representativos sin una relación directa.
- Existen muchas empresas que proveen el servicio de tokenización, pero lo hacen sin detallar la forma en la que se realiza [3]-[5].
- En 2011, el PCI publicó su guía de tokenización [6].



Arquitectura de sistema tokenizador: transacción bancaria.

Clasificación de los algoritmos tokenizadores

Clasificación del PCI [6]

- **Reversibles:** se puede regresar, a partir del token, al número de tarjeta original.
  - **Criptográficos:** cifran la tarjeta y descifran el token.
  - **No criptográficos:** utilizan una base de datos para guardar la relación entre números de tarjeta y tokens.
- **Irreversibles:** no se puede regresar al número de tarjeta a partir del token.
  - **Autenticables:** permiten validar cuando un token corresponde a un número de tarjeta dado.
  - **No autenticables:** no se puede hacer la validación anterior.

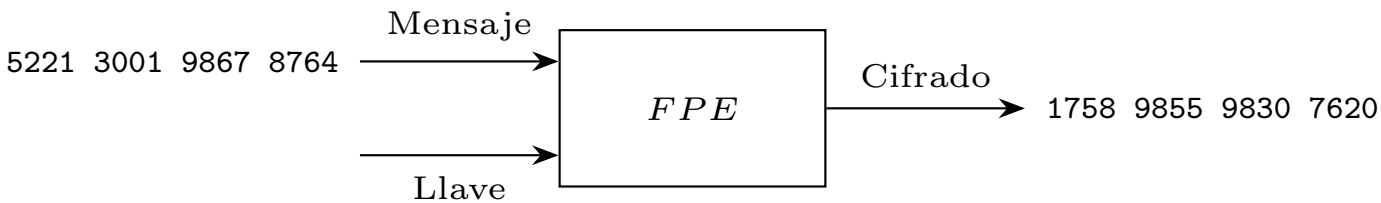
Clasificación de los algoritmos tokenizadores

Clasificación propuesta

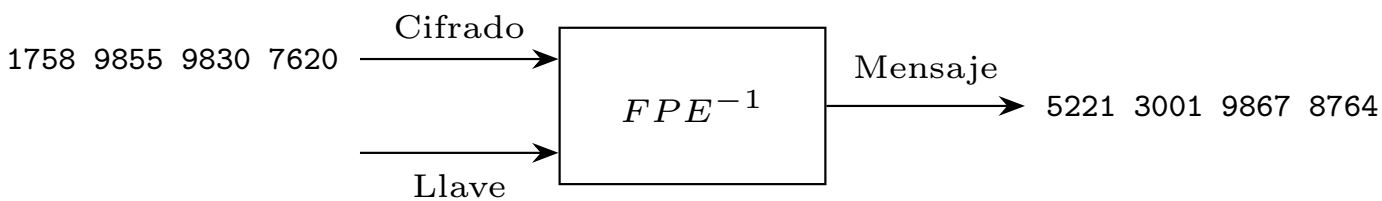
- **Criptográficos:** ocupan primitivas criptográficas en su operación.
  - **Reversibles:** cifran la tarjeta y descifran el token.
  - **Irreversibles:** requieren una base de datos para guardar la relación entre números de tarjetas y tokens.
- **No criptográficos:** no utilizan nada relacionado con la criptografía.

Métodos reversibles: FFX y BPS

- Métodos que utilizan cifrados que preservan el formato.
- Cifran la tarjeta y descifran el token.
- Se volvieron estándares en 2016 y fueron renombrados por el NIST a FF1 y FF3 respectivamente [7].
- Están basados en redes Feistel.



(a) Proceso de tokenización



(b) Proceso de detokenización

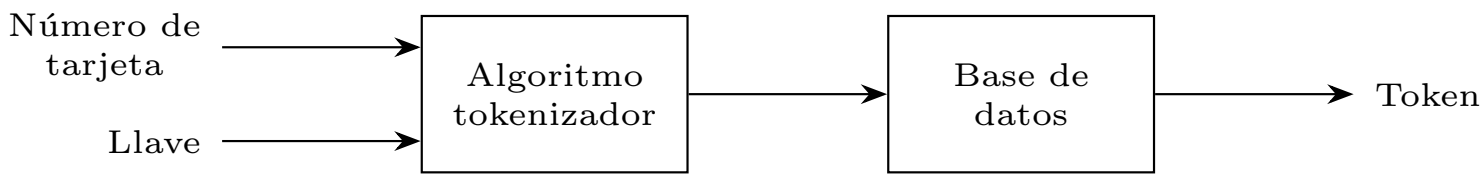
Comparativa: FFX y BPS

Características	**FFX**	**BPS**
Longitud de cadena (en caracteres)	$4 - 36$	$0 - 1.9x10^{40}$
Primitivas criptográficas	AES CBC-MAC	AES
Tamaño de llave	128 bits	128 bits
Tamaño de *tweak*	menor a 2^{64} bits	64 bits
Número de rondas	12, 24 o 28	mínimo 8 recomendadas

Características de los algoritmos tokenizadores reveribles [8], [9].

Métodos irreversibles: TKR, AHR y DRBG

- Utilizan varias primitivas criptográficas (cifrados por bloque, funciones hash, generadores pseudoaleatorios).
- Requieren guardar la relación tarjeta-token.
- Su desempeño está ligado a la base de datos.



(c) Proceso de tokenización

(d) Proceso de detokenización

Métodos irreversibles: TKR, AHR y DRBG

Características	**TKR**	**AHR**	**DRBG**
Primitivas criptográficas	Cifrado por bloque.	Cifrado por bloque y función hash.	Función hash o cifrado por bloque.
Tamaño de llave	16 bytes	32 bytes	-
¿Utiliza *tweak*?	Sí	Sí	No

Características de los algoritmos tokenizadores irreversibles [10]-[12].

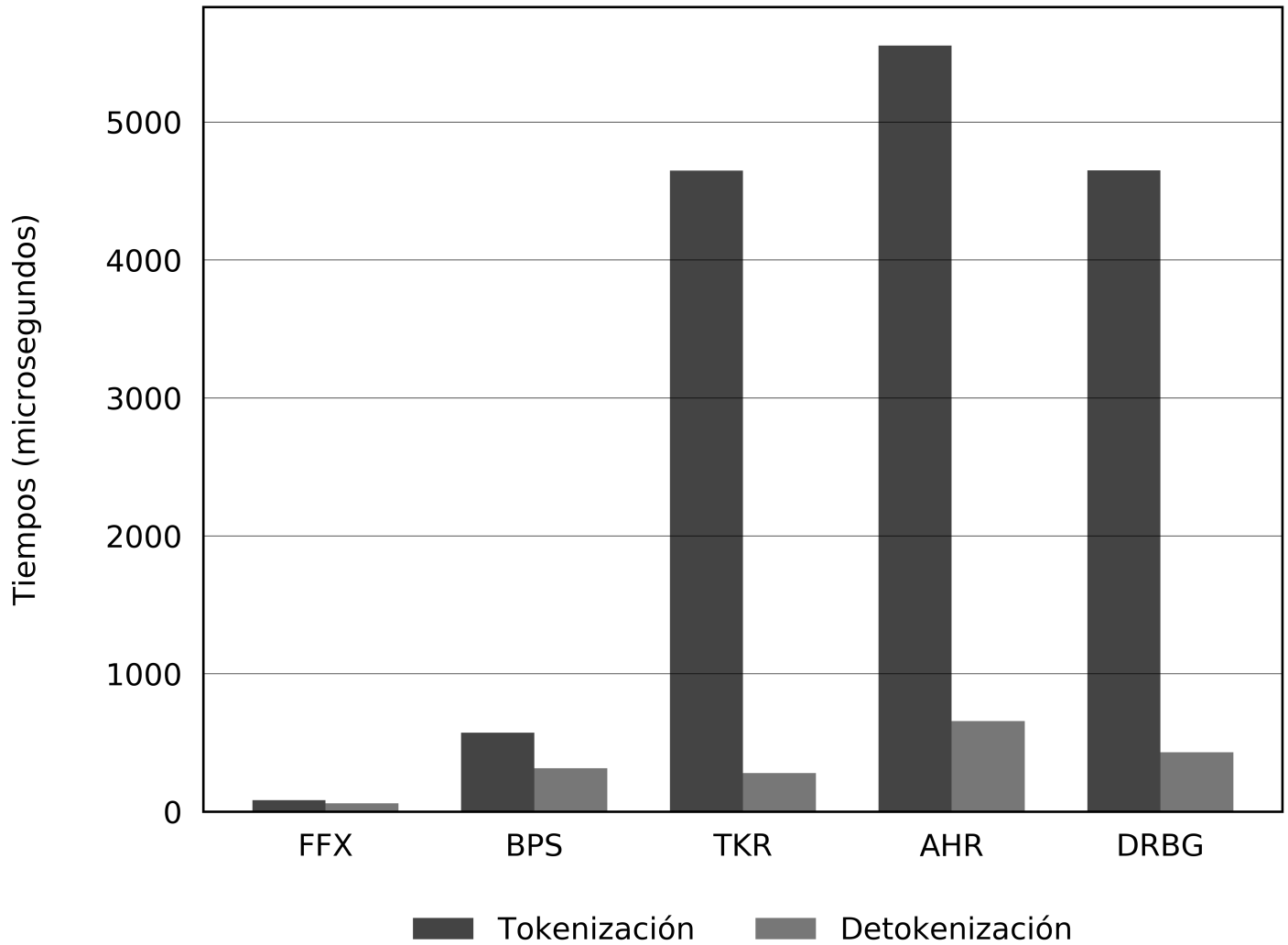
Resultados

Las pruebas de desempeño de llevaron a cabo en una computadora con las siguientes características:

- **Procesador:** Intel i5-7200U (2.5 GHz) de 4 núcleos.
- **Sistema operativo:** Arch Linux, kernel 4.18.
- **Base de datos:** MariaDB 10.1.
- **Compilador:** GCC 8.1.1.

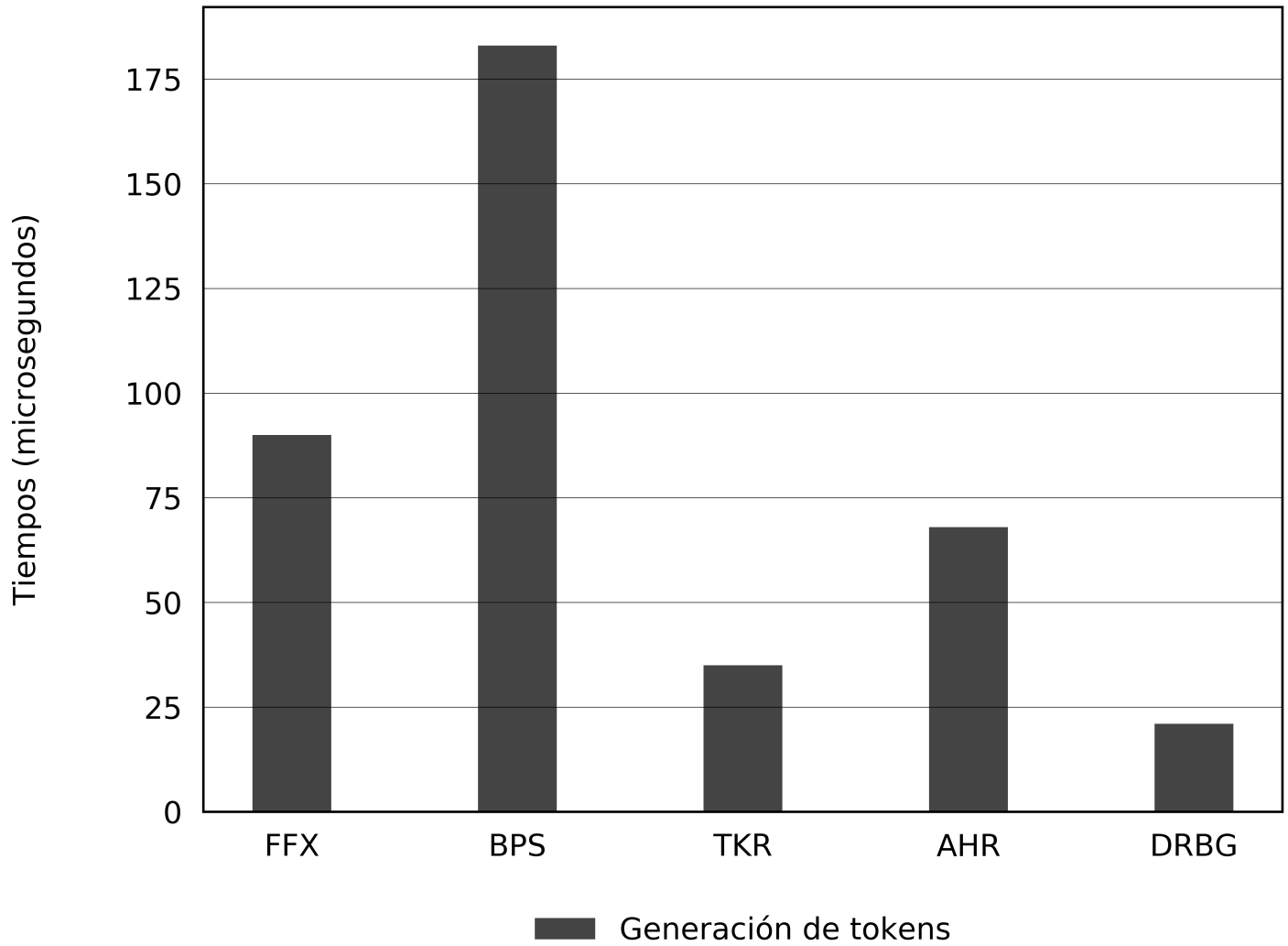
El procesador utiliza los conjuntos de instrucciones de Intel AES-NI y RD-SEED [13].

Resultados



Tokenización y detokenización.

Resultados



Generación de tokens.

Conclusiones

- La tokenización es una aplicación de la criptografía.
- La denominación *no criptográfica* del PCI es contradictoria.
- Los algoritmos reversibles son más últiles cuando se necesita tanto tokenizar como detokenizar con frecuencia.
- Los algoritmos irreversibles son más últiles cuando se requiere detokenizar con frecuencia.

Bibliografía I

[1] John S. Kiernan. *Credit Card And Debit Card Fraud Statistics*. https://wallethub.com/edu/credit-debit-card-fraud-statistics/25725/. Consultado en marzo de 2018 (vid. pág. 3).

[2] Payment Card Industry Security Standards Council. *Data Security Standard - Version 3.2*. 2016. URL: https://www.pcisecuritystandards.org/documents/pci_dss_v3-2.pdf (vid. pág. 4).

[3] Shift4 Payments. *The History of TrueTokenization*. https://www.shift4.com/dotn/4tify/trueTokenization.cfm. Consultado en agosto de 2018 (vid. págs. 6, 7).

Bibliografía II

[4] Braintree. *Tokenization Secures CC Data and Meet PCI Compliance Requirements.* https://www.braintreepayments.com/blog/using-tokenization-to-secure-credit-card-data-and-meet-pci-compliance-requirements/. Consultado en marzo de 2018 (vid. págs. 6, 7).

[5] Securosis. *Understanding and Selecting a Tokenization Solution.* https://securosis.com/assets/library/reports/Securosis_Understanding_Tokenization_V.1_.0_.pdf. Consultado en febrero de 2018 (vid. págs. 6, 7).

Bibliografía III

[6] 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 (vid. págs. 6-8).

[7] Morris Dworkin. *NIST Special Publication 800-38G - Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption*. 2016. URL: http://dx.doi.org/10.6028/NIST.SP.800-38G (vid. pág. 10).

[8] Mihir Bellare, Phillip Rogaway y Terence Spies. "The FFX Mode of Operation for Format-Preserving Encryption". Ver. 1.0. En: (2009). Presentado al NIST para estandarización (vid. pág. 11).

Bibliografía IV

[9] Eric Brier, Thomas Peyrin y Jacques Stern. "BPS: a Format-Preserving Encryption Proposal". En: (2010). Presentado al NIST para estandarización (vid. pág. 11).

[10] 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. pág. 13).

[11] 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. pág. 13).

Bibliografía V

[12] 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. pág. 13).

[13] Gael Hofemeier y Robert Chesebrough. *Introduction to Intel AES-NI and Intel Secure Key Instructions.* https://software.intel.com/sites/default/files/m/d/4/1/d/8/Introduction_to_Intel_Secure_Key_Instructions.pdf. Consultado en abril de 2018. 2014 (vid. pág. 14).

Gracias por su atención.

Un vistazo a la tokenización

Presentan

Daniel Ayala Zamorano

daz23ayala@gmail.com

Laura Natalia Borbolla Palacios

ln.borbolla.42@gmail.com

Ricardo Quezada Figueroa

qf7.ricardo@gmail.com

Sandra Díaz Santiago

sdiazs@gmail.com

Primera Reunión de Ciberseguridad para la Industria 4.0
Puebla, 14 de octubre de 2018

Escuela Superior de Cómputo
Instituto Politécnico Nacional