Aplicación del cifrado con preservación del formato para eventos de ciberseguridad

L. González-Manzano[1], J. M. de Fuentes[1], V. Gayoso Martínez[2]
[1]Computer Security Lab (COSEC). Universidad Carlos III de Madrid
[2]Instituto de Tecnologías Físicas y de la Información. Consejo Superior de Investigaciones Científicas (CSIC)
{lgmanzan, jfuentes}@inf.uc3m.es, {victor.gayoso}@iec.csic.es

***Resumen*-La compartición de información sobre eventos de ciberseguridad presenta múltiples retos. Uno de ellos es la ocultación de la información transmitida, de forma que pase inadvertida a un observador. Además de la esteganografía, se han propuesto mecanismos de cifrado que persiguen un fin similar. Uno de ellos es el cifrado con preservación del formato (*format preserving encryption*). En este artículo se describe el trabajo en curso sobre su uso en el contexto de la ciberseguridad.**

I. Introducción

El ciberespacio juega un papel fundamental en las sociedades modernas. Sin embargo, la Estrategia Nacional de Seguridad pone de manifiesto que el ciberespacio es un entorno muy accesible y de difícil control [1].

Para responder adecuadamente a estas amenazas es crítico que las organizaciones cooperen entre sí. Un pilar esencial de dicha cooperación es la compartición de información de ciberseguridad sobre vulnerabilidades, amenazas, actores, tácticas, incidentes en curso, contramedidas, etc. [2].

Este trabajo se centra en la protección de la información intercambiada acerca de eventos de ciberseguridad. Se define un *evento de seguridad* como un fallo o problema que compromete los sistemas de información, los servicios o la red, es decir, indica que una política de seguridad ha sido violada o que una salvaguarda ha fallado [7].

Un problema abierto en este entorno es que la propia existencia de la comunicación revela cierta información a un observador. Por ejemplo, si dos entidades cooperantes intercambian datos en un momento determinado, es posible que una de ellas haya sufrido un evento de seguridad. Es necesario proporcionar algún mecanismo que impida esta circunstancia.

En este artículo se describe el trabajo en curso para la aplicación de una técnica de cifrado en este contexto. Particularmente, la técnica considerada es el cifrado con preservación del formato (en adelante FPE, del inglés *Format Preserving Encryption*) [3].

II. Nombrado de eventos de ciberseguridad

El foco de este trabajo se sitúa en posibilitar la compartición de forma inadvertida para un tercero. Como paso previo, es necesario explorar las herramientas existentes para una compartición efectiva.

En los últimos tiempos, se han propuesto numerosos mecanismos para intercambiar esta información [4]. Además, es imprescindible que todas las partes nombren los eventos de seguridad de forma unívoca. En concreto se han propuesto algunos formatos unificados para nombrar los eventos de ciberseguridad. En esta sección se revisan brevemente dos formatos representativos. El uso de formatos bien definidos es la base fundamental para la aplicación del FPE.

A. *Recomendación ITU-T X.1520*

La recomendación X.1520 de la ITU-T persigue proporcionar una *forma estructurada de intercambio global de información públicamente conocida sobre vulnerabilidades y exposiciones maduras* [5]. Para ello, determina el correcto uso de los CVE (*Common Vulnerabilities and Exposures*). Dichos CVE aglutinan las diferentes vulnerabilidades de diversos programas y plataformas.

B. *STIX*

STIX (*Structured Threat Information Expression*) es una iniciativa de la corporación MITRE para estructurar esta información [6]. En dicho lenguaje se definen numerosos conceptos: ciber-observable, indicadores, táctica/técnica/procedimiento de ataque, plataforma objetivo, tipo de atacante, etc. Por ejemplo, en la Figura 1 se muestra parte del indicador de un ataque de *phishing*. Se describe el título, el tipo de ataque, cómo ha ocurrido y la fecha del mismo.

```
<stix:Indicators>
  <stix:Indicator xsi:type="indicator:IndicatorType"
    id="example:Indicator-19e5d914-cc0e-478f-a523-b099a34383f7"
    timestamp="2014-02-20T09:00:00.000000Z">
    <indicator:Title>"US-China" Phishing Indicator</indicator:Title>
    <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">Malicious
      E-mail</indicator:Type>
    <indicator:Description>This is a cyber threat indicator for instances of "US-China"
      phishing attempts.</indicator:Description>
    <indicator:Valid_Time_Position>
      <indicator:Start_Time>2012-12-01T09:30:47Z</indicator:Start_Time>
      <indicator:End_Time>2013-02-01T09:30:47Z</indicator:End_Time>
    </indicator:Valid_Time_Position>
```

Fig. 1 Indicador (porción) de un ataque de phising utilizando STIX [11].

III. Cifrado con preservación del formato (FPE)

Los sistemas de cifrado FPE se basan en que el formato del texto cifrado sea equivalente al del texto en claro. Este término fue acuñado por Terence Spies [12, 13] proponiendo distintas aplicaciones, entre las que destacan el cifrado de números de la seguridad social o números de tarjetas de crédito (Figura 2).

FPE hace uso de una clave simétrica K utilizada para cifrar un texto X de un formato N descrito en el conjunto finito $\aleph_N$, donde $\aleph_N$ se corresponde con todos los posibles espacios de los

mensajes, es decir, con los posibles valores de textos en claro y cifrados. Para cifrar $X \in \aleph_N$ se realiza la llamada a un algoritmo de cifrado ε proporcionando la clave K y el texto X. Esto producirá un texto cifrado $Y = E_K(X)$ manteniendo que $Y \in \aleph_N$. El proceso de descifrado, $D_K(Y)$, proporcionará nuevamente el texto X. Además, $E_K(\cdot)$ ha de corresponderse con una permutación aleatoria en el espacio de mensajes $\aleph_N$.

En base a la descripción realizada, un buen sistema FPE es aquel en el que un adversario no es capaz de distinguir un texto cifrado con una permutación $E_K(\cdot)$ y ese mismo texto cifrado con una permutación $\pi_K(\cdot)$.

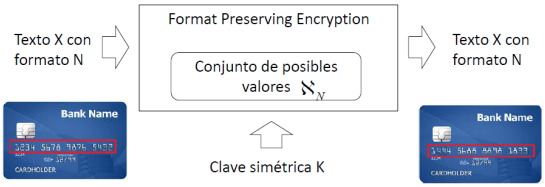

Fig. 2. Esquema de cifrado con preservación de formato

A. *Aproximación propuesta para eventos de ciberseguridad*

Para utilizar FPE en relación con eventos de seguridad se pueden distinguir los siguientes pasos:

1. Establecer qué formato se va a considerar.
2. Determinar el espacio de mensajes $\aleph_N$. ¿Se cifran los datos del tipo de ataque, acciones a emprender, etc. o sólo un dato concreto, por ejemplo el elemento observado?
3. Realizar análisis de los sistemas FPE existentes y escoger el más apropiado en su caso, modificándolo o desarrollando uno que satisfaga las necesidades. Asumiendo que $\aleph_N$ es el espacio de mensajes compuesto por los caracteres del alfabeto y los números, se podría estudiar la posibilidad de utilizar el algoritmo presentado en [14]. Así se conseguiría el cifrado de caracteres en base a su código binario mediante un cifrador de bloque basado en una estructura Feistel.

En relación con el ejemplo presentado en la Figura 1, un cifrado FPE se muestra en la Figura 3 (enmarcado en rojo). En este caso, por ejemplo, el cifrado de "Malicious E-mail" se corresponde con "Malicious Docume", así como la fecha de comienzo, "2012-12-01", que puede cifrarse como "2013-02-12".

```
<stix:Indicators>
  <stix:Indicator xsi:type="indicator:IndicatorType"
    id="example:Indicator-19e5d914-cc0e-478f-a523-b099a34383f7"
    timestamp="2014-02-20T09:00:00.000000Z">
    <indicator:Title"Luxembo" Networks Indicator</indicator:Title>
    <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">Malicious
     Docume</indicator:Type>
    <indicator:Description>This is a cyber threat indicator for instances of "Luxembo"
    Networks attempts.</indicator:Description>
    <indicator:Valid_Time_Position>
      <indicator:Start_Time>2013-02-12T23:18:95Z</indicator:Start_Time>
      <indicator:End_Time>2015-02-20T10:12:17Z</indicator:Start_Time>
    </indicator:Valid_Time_Position>
```

Fig 3. Cifrado FPE de indicador de *phishing*

IV. Líneas de trabajo en curso

En esta línea de trabajo existen numerosos aspectos que merecen atención investigadora. En primer lugar, el desarrollo de técnicas FPE ligeras permitiría un intercambio ágil de información. Este tipo de primitivas podría inspirarse en las de la criptografía ligera (*lightweight cryptography*) [9].

Por otro lado, una cuestión importante es aplicar esta técnica de forma que los resultados pudiesen correlarse. La correlación permite conocer el ataque que se padece o el grado de avance del mismo [10]. Esto tendría un elevado potencial disuasorio para el observador, quien no podría establecer si está visualizando los eventos reales o el resultado de cifrarlos.

La aleatoriedad de los resultados es también crítica en este entorno. Los eventos de seguridad pueden repetirse con gran frecuencia, lo que podría dar lugar a ataques de texto claro escogido que permitiesen revelar la clave. Se necesita, por tanto, que los resultados sean impredecibles.

V. Conclusiones

La compartición de información sobre eventos de ciberseguridad ha recibido gran atención por parte de la comunidad investigadora en los últimos tiempos. Entre los numerosos retos que ésta plantea, este artículo ha presentado el trabajo en curso para la protección de dicho intercambio. Particularmente, se propone desarrollar una aproximación basada en el cifrado con preservación del formato (FPE). Gracias a FPE, el observador no es capaz de distinguir si visualiza el mensaje en claro o su resultado cifrado, en tanto que ambos mantienen la misma estructura.

Agradecimientos

Este trabajo ha sido financiado a través del proyecto CIBERDINE (S2013/ICE-3095), otorgado por la Comunidad Autónoma de Madrid, y del proyecto SPINY (TIN2013-46469-R), otorgado por el MINECO.

Referencias

[1] España, "Estrategia Nacional de Seguridad", 2013.
[2] E. Gal-or y A. Ghose, "The economic incentives for sharing security information". *Information Systems Research*, 2005, vol. 16, p. 186-208.
[3] M. Bellare, *et al.*, "Format-preserving encryption". https://eprint.iacr.org/2009/251.pdf (último acceso Julio 2015).
[4] L. Dandurand, "Cyber Defense Data Exchange and Collaboration Infrastructure (CDXI)". *ITU-T Workshop*, 2010
[5] ITU-T, "X.1520: Vulnerabilidades y estados comunes", 2014.
[6] S. Barnum, "Standardizing Cyber Threat Intelligence Information with the STIX", disponible en http://stixproject.github.io/getting-started/whitepaper/ (último acceso Julio 2015).
[7] ISO, ISO27001-Glosario, disponibles en http://www.praxiom.com/iso-27001-definitions.htm (último acceso Julio 2015).
[8] M. Brightwell and H. Smith. "Using datatype-preserving encryption to enhance data warehouse security". 20th NISSC Proc., pp. 141–149, 1997
[9] T. Eisenbarth, "A survey of lightweight-cryptography implementations". *IEEE Design & Test of Computers*, 2007, p. 522-533.
[10] F. Cuppens, A. Miege, "Alert correlation in a cooperative intrusion detection framework.". *Security and Privacy, 2002. Proceedings. 2002 IEEE Symposium on*. IEEE, 2002. p. 202-215.
[11] Ejemplos de STIX, disponible en: http://stix.mitre.org/about/example_phishing.html (últ. acceso Julio 2015)
[12] Spies, Terence. "Format preserving encryption." Unpublished white paper, www. voltage. com Database and Network Journal. 2008.
[13] Black, John, and Phillip Rogaway. "Ciphers with arbitrary finite domains." Topics in Cryptology—CT-RSA 2002. Springer Berlin Heidelberg, 2002. 114-130
[14] Li, Min, et al. "Format-preserving encryption for character data." Journal of Networks 7.8 (2012): 1239-1244.