Manga Guide To Cryptography

Por:

Enrique Ulises Báez Gómez Tagle

Trabajo presentado para
la asignatura Criptografía y Seguridad en Redes

Universidad Panmericana
Facultad de Ingeniería
Ciudad de México, México

16 de noviembre de 2024

Manga Guide To Cryptography

Enrique Ulises Báez Gómez Tagle

16 de noviembre de 2024

Índice

1. Objetivo **2**

2. Introducción **2**

3. Desarrollo **2**
3.1. Fundamentos de Encriptación 2
3.1.1. El Cifrado 2
3.1.2. Modelo de Shannon 3
3.1.3. Métodos Clásicos 3
3.1.4. Seguridad 3
3.2. Algoritmos de Llave Simétrica 3
3.2.1. El ABC Digital: Bits y Operaciones 3
3.2.2. Familias de Cifrado 4
3.2.3. DES 4
3.2.4. Los nuevos algoritmos: 3-DES y AES 4
3.3. criptografía asimétrica 4
3.4. Funciones Unidireccionales 4
3.5. Números Primos y Factorización 4
3.6. Operaciones de Módulo 5
3.7. Teoremas de Fermat y Euler 5
3.8. Aplicaciones Prácticas de la Encriptación 5
3.8.1. Funciones Hash 5
3.8.2. Encriptación Híbrida 5
3.8.3. PKI (Infraestructura de Clave Pública) 5

4. Conclusiones **6**

5. Bibliografía **6**

1. Objetivo

"The Manga Guide to Cryptography"nos enseña varios de los temas fundamentales de la seguridad de datos, algo crítico en el área de la inteligencia de datos (donde nosotros nos desempeñamos actualmente). Este libro toca conceptos como los algoritmos de cifrado simétrico o asimétrico, que por ejemplo, son cruciales para proteger la integridad de los datos cuando hacemos procesamiento y análisis en entornos de big data.

Otros temas que se tocan son la implementación de funciones hash y técnicas de autenticación, que también son indispensables en el pipeline de datos actual, o por ejemplo, los conceptos de PKI también son súper relevantes en los entornos de datos distribuidos y para la implementación de data lakes seguros.

Para profesionales de datos como nosotros, es importantísimo comprender estos fundamentos de criptografía, pues la protección de datos sensibles y el cumplimiento de regulaciones de privacidad son inherentes al ciclo de vida del análisis de datos. Este libro nos ayudará a comprender estos temas de manera efectiva y a aplicarlos en nuestros proyectos.

2. Introducción

The Manga Guide es una serie súper interesante que nos acerca a conceptos técnicos usando el formato manga. En el caso de la criptografía, esto es especialmente útil para nosotros como científicos de datos, ya que nos permite entender conceptos complejos de seguridad a través de una historia entretenida con imagenes y hasta elementos de misterio.

Aquí, seguimos a Ruka Meguro, al Inspector Jun Meguro y a Rio Yoneda, una reportera. Juntos intentan resolver el misterio de Ms. Cypher, una "ladrona fantasma". A través de esta trama, vamos aprendiendo los fundamentos de la criptografía que aplicamos en nuestro trabajo diario con datos.

Es increíble cómo la criptografía está presente en casi todo lo que hacemos en el análisis de datos, desde la protección de bases de datos hasta la seguridad en nuestros pipelines de procesamiento. Este libro nos ayuda a entender estos conceptos de forma práctica y amena. Si trabajas con datos, te recomiendo muchísimo leerlo completo, ya que explica conceptos fundamentales que necesitamos manejar en nuestro día a día como profesionales de datos.

3. Desarrollo

3.1. Fundamentos de Encriptación

3.1.1. El Cifrado

El cifrado es la técnica que usamos para proteger la información sensible, convirtiendo información legible en código secreto que solo pueden descifrar las personas autorizadas.

3.1.2. Modelo de Shannon

Shannon revolucionó la criptografía cuando estableció que todo proceso de cifrado requiere un texto original, un método de encriptación (E) y una llave (K), siendo la desencriptación el proceso inverso usando un algoritmo de descifrado (D).

3.1.3. Métodos Clásicos

- **Cifrado César:** Simplemente consiste en avanzar cada letra un número fijo de espacios en el abecedario.
- **Cifrado de Sustitución:** Hay que reemplazar las letras del mensaje usando una guía que nos dice qué letra va por cuál.
- **Cifrado Polialfabético:** Es más complejo porque usa diferentes reglas de sustitución que van cambiando según avanzamos en el texto.
- **Cifrado de Transposición:** Aquí, las letras son las mismas, solo que las mezclamos siguiendo cierta regla establecida.

3.1.4. Seguridad

Para poder saber qué tan seguro puede ser un método de cifrado, hay que contemplar dos cosas principales:

- Cuántas llaves diferentes podemos usar (espacio de claves)
- Qué tan resistente es contra alguien que intente analizar patrones en el texto

Cifrados como el de Vernam son "matemáticamente imposibles"de romper porque usan llaves aleatorias tan larga como el mensaje mismo. Pero siendo realistas, esto no es muy práctico.

Por eso en el mundo real usamos cifrados que son "suficientemente seguros", no 100 % seguros pero usan mucho tiempo computacional o recursos computacionales que sería caro usarlos.

3.2. Algoritmos de Llave Simétrica

Literal como si tú y tu amigo tienen la misma llave para un candado. La ventaja de la criptografía simétrica es que es rápida y muy buena al momento de manejar muchos datos, aunque mucho ojo a quién le compartimos la llave.

3.2.1. El ABC Digital: Bits y Operaciones

Las computadoras hablan en unos y ceros (bits), y juntan estos en grupos de 8 (bytes). Para cifrar, usan trucos matemáticos como el XOR, así podemos recuperar cualquier parte del proceso si tenemos las otras dos.

3.2.2. Familias de Cifrado

- **Cifrados de Flujo** Estos cifran la información poco a poco, como una corriente continua. Usan una receta secreta"muy larga para generar números que parecen aleatorios. RC4 y SEAL son algunos ejemplos famosos.
- **Cifrados por Bloques** Funcionan como un rompecabezas: dividen el mensaje en piezas iguales y las cifran una por una. Pueden trabajar con piezas de diferentes tamaños y tienen distintas formas de hacerlo (como ECB y CBC).

3.2.3. DES

DES fue el pionero en la encriptación comercial. Funciona dividiendo el mensaje en bloques y aplicando varias "vueltas"de cifrado. Aunque en su momento fue revolucionario, hoy en día su llave es demasiado pequeña para considerarse segura.

3.2.4. Los nuevos algoritmos: 3-DES y AES

- **3-DES:** Como aplicar DES, pero tres veces seguidas, usando tres llaves diferentes.
- **AES:** Es como un laberinto super complejo que mezcla y revuelve la información de formas muy sofisticadas, y usa una red de sustitución-permutación (SPN).

3.3. criptografía asimétrica

Es como si fuera un candado mágico que cualquiera puede cerrar, pero solo tú puedes abrir. La criptografía asimétrica: usa una llave pública (para que otros te envíen mensajes) y una privada (para que solo tú los leas).

3.4. Funciones Unidireccionales

Tal cual, un tobogán matemático: fácil deslizarse hacia abajo, imposible subir. Es fácil multiplicar números grandes pero muy difícil encontrar sus factores.

3.5. Números Primos y Factorización

Los números primos son escenciales en de la criptografía, principalmente para encriptar nuestra llave pública, y hay varios algoritmos que basan su seguridad en números compuestos grandísimos que son difíciles de factorizar en primos.

3.6. Operaciones de Módulo

Se usan operaciones de módulo para las 4 operaciones básicas: suma, resta, multiplicación y divisón que igual se usa para encriptar por ejemplo llaves públicas. Y estas operaciones solo se pueden hacer a ciertos números con características específicas.

3.7. Teoremas de Fermat y Euler

- El Teorema de Fermat describe el comportamiento de los números primos cuando se realizan operaciones de potenciación modular.
- El Teorema de Euler generaliza este concepto para cualquier número compuesto, introduciendo la función totient que cuenta números coprimos.

3.8. Aplicaciones Prácticas de la Encriptación

La encriptación va más allá de proteger la confidencialidad de datos, abarcando la integridad, autenticidad y no repudio de la información. Algunas de las que más me llamaron la atención por el medio en el que me desempeño actualmente fueron:

3.8.1. Funciones Hash

"Huellas digitales"única para cada mensaje o los datos. Sirven para: - Deduplicación de datos - Anonimización de información sensible (como PII) - Indexación y búsqueda rápida en grandes conjuntos de datos - Verificación de integridad en pipelines de datos

3.8.2. Encriptación Híbrida

Combina la velocidad de la encriptación simétrica para los datos y la seguridad de la asimétrica para el intercambio de claves. La usamos para: - Almacenamiento seguro en data lakes - Transferencia de grandes volúmenes de datos - Protección de modelos de ML en producción - Securización de APIs de servicios de datos

3.8.3. PKI (Infraestructura de Clave Pública)

Es todo el sistema que permite que las firmas y certificados digitales funcionen de forma confiable. Se aplica en: - Autenticación segura en servicios cloud de ML - Certificación de origen de datos para modelos - Gestión de accesos a recursos compartidos - Verificación de integridad en pipelines de MLOps

4. Conclusiones

A lo largo de esta lectura pude comprender los fundamentos de la criptografía y e identificar su aplicación en el análisis de datos para entenderlos mejor. Hoy en día las técnicas de encriptación son fundamentales para el procesamiento seguro de información. Y con lo aprendido en esta lectura, vemos que es posible balancear la seguridad con el rendimiento requerido en entornos de datos modernos.

En las funciones hash y la encriptación híbrida encontré nuevas herramientas para proteger y diseñar data lakes y pipelines de datos, y también me gustaría probar la infraestructura PKI para intentar facilitar el despliegue seguro de modelos de machine learning en entornos de producción.

5. Bibliografía

Mitani, M., Sato, S., & Trend-Pro Co., Ltd. (2018). The manga guide to cryptography. No Starch Press.