TECNICATURA UNIVERSITARIA
EN PROGRAMACIÓN
A DISTANCIA

TP 8 - Unidad 8
Seguridad y Privacidad en Bases de Datos

Nombre: Farias, Gustavo
Comisión: M2025-13
Matrícula: 101662
Repositorio GitHub:
https://github.com/Lucenear/UTN-TUPaD-TPs/tree/main/Bases%20de%20Datos/Bases%20de%20Datos%20I

1. Explorando el Cifrado Homomórfico para la Protección de Datos en Uso

¿Qué es el cifrado homomórfico? (Definición y concepto básico)

El cifrado homomórfico es una forma avanzada de criptografía que permite realizar operaciones matemáticas (como sumas o multiplicaciones) directamente sobre datos cifrados, sin necesidad de descifrarlos previamente. El resultado de estas operaciones, que también permanece cifrado, al ser descifrado es idéntico al resultado que se habría obtenido si las operaciones se hubieran realizado sobre los datos originales en texto plano.

Cuál es su principio fundamental y cómo se diferencia de otras técnicas de cifrado más tradicionales?

Principio Fundamental: Se basa en la propiedad de la "homomorfía", que en matemáticas significa que la estructura de una operación se preserva a través de una transformación. En criptografía, esto se traduce en que las operaciones realizadas sobre el texto cifrado se corresponden directamente con las mismas operaciones aplicadas al texto plano subyacente (por ejemplo, Cifrado(A) + Cifrado(B) = Cifrado(A+B)).

Diferencias con el Cifrado Tradicional:

- Cifrado Tradicional: Protege los datos en reposo (almacenados) y en tránsito (en la red). Sin embargo, para procesar los datos (datos en uso), el sistema debe descifrarlos primero, creando una ventana de vulnerabilidad donde la información sensible está expuesta en la memoria del servidor.
- Cifrado Homomórfico: Extiende la protección a los datos en uso. Los datos nunca se descifran durante el procesamiento, eliminando esa ventana de vulnerabilidad y permitiendo que terceros (como un proveedor de nube) realicen cálculos sin acceder nunca a la información original.

Identifique al menos dos ventajas clave y dos desafíos o limitaciones para su implementación

Ventajas:

1.

- Privacidad durante el procesamiento: Permite externalizar cálculos a entornos no confiables (como la nube pública) sin comprometer la confidencialidad de los datos sensibles.
- Habilitación de Colaboración Segura: Facilita el análisis de datos confidenciales de múltiples fuentes (ej., entre hospitales para investigación médica) sin necesidad de compartir los datos crudos.

Desafíos/Limitaciones:

1.

- Baja Eficiencia Computacional: Las operaciones sobre datos cifrados son significativamente más lentas y consumen muchos más recursos (CPU, memoria) que las operaciones equivalentes sobre datos en texto plano, lo que lo hace poco práctico para aplicaciones que requieren respuestas en tiempo real.

- Complejidad de Implementación: La tecnología es aún emergente y compleja, con falta de estandarización y de integración directa en los sistemas de bases de datos convencionales, requiriendo conocimientos criptográficos especializados.

Mencione al menos dos ejemplos concretos de aplicaciones potenciales

1) Análisis de Datos Médicos en la Nube: Un hospital podría subir registros de pacientes cifrados homomórficamente a un servidor en la nube para que se ejecuten algoritmos de análisis estadístico. El proveedor de la nube realizará los cálculos sin poder ver la información médica confidencial, y solo el hospital podría descifrar el resultado final.
2) Aprendizaje Automático (Machine Learning) con Privacidad Preservada: Se podría entrenar un modelo de IA utilizando datos de entrenamiento cifrados provenientes de múltiples entidades (ej., diferentes bancos). El modelo aprendería de los patrones generales sin que ninguna de las partes ni el servidor de entrenamiento tuvieran acceso a los datos sensibles individuales de los demás.

2. Técnicas de Auditoría y Monitoreo en Bases de Datos Relacionales

¿Cuál es el propósito principal de la auditoría y el monitoreo en la seguridad de bases de datos?

El propósito principal es triple:

- Detección de Amenazas: Identificar comportamientos anómalos o maliciosos en tiempo real o de manera retrospectiva.
- Análisis Forense: En caso de una brecha de seguridad, los registros de auditoría son cruciales para determinar qué sucedió, cuándo, cómo y quién fue el responsable, permitiendo entender el alcance del incidente.
- Cumplimiento Legal: Proveer la evidencia necesaria para demostrar que se están implementando controles de seguridad adecuados, tal como lo exigen las regulaciones de privacidad de datos.

Describa al menos tres tipos de actividades o eventos específicos que deben ser monitoreados

- Intentos de Inicio de Sesión Fallidos: Un número elevado de intentos fallidos, especialmente desde múltiples direcciones IP, puede indicar un ataque de fuerza bruta para adivinar credenciales.
- Cambios en los Permisos de Usuario (Escalación de Privilegios): Cualquier modificación en los privilegios de un usuario debe ser monitoreada, ya que podría ser una señal de que un atacante está intentando obtener mayores accesos dentro del sistema.
- Accesos a Tablas con Datos Sensibles: El acceso (lectura o modificación) a tablas que contienen información personal, financiera o de salud debe ser rastreado para detectar consultas inusuales o no autorizadas por parte de usuarios que normalmente no interactúan con esos datos.

Explique cómo las herramientas SIEM pueden mejorar la detección de incidentes

Un Sistema de Gestión de Información y Eventos de Seguridad (SIEM) mejora la detección de incidentes al:

- Centralizar y Correlacionar Eventos: Recoge logs de auditoría de la base de datos y de otras fuentes (sistemas operativos, firewalls, aplicaciones). En lugar de analizar logs aislados, el SIEM puede identificar patrones complejos. Por ejemplo, puede relacionar un intento de inicio de sesión fallido desde una IP externa con un acceso posterior a una tabla crítica desde esa misma IP, lo que un administrador humano podría pasar por alto al revisar logs por separado.

- Generar Alertas en Tiempo Real: Al detectar patrones que coinciden con una regla de seguridad predefinida (ej., 10 intentos de login fallidos en 2 minutos), el SIEM genera una alerta inmediata para que el equipo de seguridad investigue y responda rápidamente.

- Automatizar el Análisis: Utiliza análisis basado en reglas y machine learning para reducir el "ruido" en los logs y destacar los eventos verdaderamente sospechosos.

Argumente la importancia de la auditoría y el monitoreo para el cumplimiento legal y normativo

La auditoría y el monitoreo no son solo buenas prácticas técnicas, sino un requisito explícito o implícito en la mayoría de las normativas de protección de datos. Por ejemplo, la Ley 25.326 (Argentina) y GDPR (UE), en donde ambas establecen el principio de "seguridad", obligando a los responsables de los datos a implementar medidas técnicas y organizativas apropiadas para garantizar un nivel de seguridad adecuado. Un programa robusto de auditoría y monitoreo es una de estas medidas clave. Además, los registros de auditoría sirven como evidencia tangible para demostrar a los organismos reguladores que la organización está vigilando activamente el acceso y uso de los datos personales, facilitando el cumplimiento de los requisitos de rendición de cuentas ("accountability"). Sin estos registros, sería casi imposible probar que se tomaron las medidas necesarias para proteger la información, arriesgándose a multas cuantiosas.

3. Estrategias de Respaldo y Recuperación para la Resiliencia de Bases de Datos

Defina la importancia fundamental de las estrategias de respaldo y recuperación

Las estrategias de respaldo y recuperación son la última línea de defensa contra la pérdida de datos y un pilar fundamental para la continuidad del negocio. Su importancia radica en garantizar la disponibilidad e integridad de la información ante una amplia gama de incidentes, desde fallos de hardware y error humanos hasta ataques de ransomware y desastres naturales. Sin una estrategia sólida, una organización podría sufrir la interrupción de sus operaciones, pérdidas financieras irreparables y un daño severo a su reputación.
Describa detalladamente los siguientes tipos de respaldo

- Respaldo Completo (Full Backup): Realiza una copia de todos los datos de la base de datos en un momento específico.

    - Cuándo se aplica: Suele ser la base de cualquier estrategia, realizada de forma periódica (ej., semanalmente). Es fundamental para tener un punto de restauración conocido y consistente.
    - Implicaciones para la recuperación: Es el más rápido y simple de restaurar, ya que solo se necesita un conjunto de archivos. Sin embargo, es el que más tiempo lleva realizar y más espacio de almacenamiento consume.

- Respaldo Diferencial (Differential Backup): Copia solo los datos que han cambiado desde el último respaldo completo.

    - Cuándo se aplica: Se realiza entre respaldos completos (ej., diariamente) para capturar los cambios recientes de manera más eficiente que un respaldo completo.
    - Implicaciones para la recuperación: La restauración requiere dos elementos: el último respaldo completo y el último respaldo diferencial. Es más lento de restaurar que un completo, pero más rápido que un incremental.

- Respaldo Incremental (Incremental Backup): Copia solo los datos que han cambiado desde el último respaldo de cualquier tipo (ya sea completo, diferencial o incremental).

    - Cuándo se aplica: Es el más eficiente en términos de tiempo y espacio, por lo que se puede realizar con mucha frecuencia (ej., cada pocas horas).
    - Implicaciones para la recuperación: La restauración es la más compleja y lenta, ya que requiere la cadena completa: el último respaldo completo y todos los respaldos incrementales subsiguientes.

- Respaldo del Registro de Transacciones (Transaction Log Backup): Captura todas las transacciones registradas desde el último respaldo de registro.

    - Cuándo se aplica: En bases de datos que operan en modelo de recuperación completa. Se realiza con mucha frecuencia (ej., cada 15 minutos o menos).
    - Implicaciones para la recuperación: Permite la Recuperación hasta un Punto en el Tiempo (PITR), minimizando la pérdida de datos al máximo. Es crucial para restaurar la base de datos a un estado consistente justo antes de un fallo.

Explique la "Regla 3-2-1" para respaldos y justifique por qué cada uno de sus componentes es vital

La Regla 3-2-1 es una estrategia fundamental para garantizar la redundancia y resiliencia de los datos. Establece:

- 3 Copias de los Datos: El dato original más al menos dos respaldos.
    - Justificación: Una sola copia es un punto único de fallo. Con tres copias, se tiene redundancia. Si una falla (por ejemplo, el disco original se corrompe), quedan otras dos.
- 2 Tipos de Medios de Almacenamiento Diferentes: Almacenar las copias en al menos dos tipos de soportes físicos diferentes.

    - Justificación: Protege contra fallos específicos de una tecnología. Por ejemplo, si todas las copias están en discos duros, un fallo eléctrico podría dañarlas todas. Al tener una en discos duros y otra en cintas o en un servicio de nube (que usa tecnología subyacente diferente), se mitiga este riesgo.
- 1 Copia Almacenada Fuera del Sitio (Off-site): Mantener al menos una copia de seguridad en una ubicación física separada.
    - Justificación: Protege contra desastres locales que podrían destruir todas las copias en una ubicación, como incendios, inundaciones, robos o sabotajes.

Discuta la relación entre las estrategias de respaldo y la definición del RPO y RTO

El RPO y el RTO son métricas críticas que definen los requisitos de una estrategia de respaldo y recuperación, y la estrategia, a su vez, está diseñada para cumplirlos.

- RPO (Objetivo de Punto de Recuperación): Define la cantidad máxima de datos que la organización está dispuesta a perder. Se mide en tiempo.
    - Relación con las Estrategias de Respaldo: El RPO determina directamente la frecuencia con la que se deben realizar los respaldos. Un RPO de 1 hora exige respaldos (especialmente de transacciones) al menos cada hora. Un RPO de 24 horas permite respaldos diarios.

- RTO (Objetivo de Tiempo de Recuperación): Define el tiempo máximo tolerable de inactividad del sistema después de un incidente.
    - Relación con las Estrategias de Respaldo: El RTO determina la velocidad con la que se debe poder restaurar el sistema. Un RTO corto (ej., 1 hora) exige estrategias de restauración rápidas, lo que influye en el tipo de respaldo utilizado (un respaldo completo es más rápido de restaurar que una cadena de incrementales) y en la infraestructura preparada para la recuperación (ej., servidores en espera).

Importancia de probar y documentar los procedimientos de recuperación

Un respaldo sin una restauración probada no es más que una copia de dudosa confiabilidad. Probar periódicamente los procedimientos de recuperación es esencial para:
- Verificar la integridad de los respaldos.
- Garantizar que el RTO se puede cumplir en la práctica.
- Capacitar al personal y refinar el proceso.

Documentar el procedimiento asegura que cualquier persona autorizada pueda ejecutarlo de manera eficiente y correcta bajo presión, minimizando errores y el tiempo de inactividad durante un desastre real.

Fuentes:

- https://servicios.infoleg.gob.ar/infolegInternet/anexos/60000-64999/64790/texact.htm
- https://dev.mysql.com/doc/refman/8.0/en/security.html
- https://learn.microsoft.com/en-us/sql/relational-databases/security/security-center-for-sql-server-database-engine-and-azure-sql-database?view=sql-server-ver17
- https://owasp.org/www-project-top-ten/
- https://www.veeam.com/blog/321-backup-rule.html