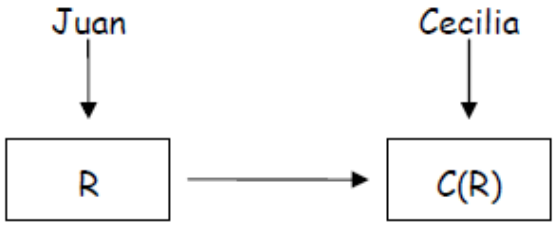
LA SEGURIDAD MULTINIVEL

Lo que se persigue con este tipo de técnicas es resguardar los caminos que toman los datos con el fin de evitar que lleguen a las manos de personas no autorizadas. El problema de control de flujos se ilustra en la figura siguiente:



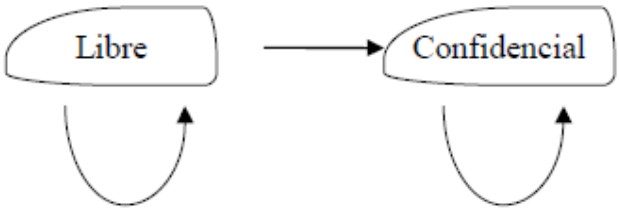
Cecilia y Juan son usuarios de un SGBD. Cecilia no tiene el derecho de acceso a la relación R, entonces, solicita a Juan, que si tiene derecho, le brinde una copia C(R) de esta relación.

En general, se puede decir que un flujo existe entre un objeto X y un objeto Y cuando un programa lee X y hace escrituras sobre Y.

Un ejemplo clásico de un flujo es el copiar un archivo X en el archivo Y.

La mayoría de los métodos de control de flujos utilizan niveles de seguridad. Así, los objetos de la base de datos se dividen en clases y cada una de ellas cuenta con un grado de seguridad. Según la transferencia de información entre un emisor y un receptor es posible sólo si el nivel de seguridad del receptor es al menos tan privilegiado como el del emisor.

El protocolo más sencillo de control de flujos comprende dos niveles: libre y confidencial. En este caso se permite cualquier flujo de datos excepto aquellos que van del nivel confidencial al nivel libre, según se aprecia en la figura:



Otros sistemas más complejos pueden requerir varios niveles, de seguridad y por ende utilizar protocolos más complejos. Así, cada nivel de seguridad se puede representar por medio de un par (nivel

de autorización, conjunto de categorías). Se pueden establecer al menos cuatro niveles de autorización:

- No clasificado
- Confidencial
- Secreto
- Muy secreto

Para estos niveles de autorización se puede establecer el siguiente orden:

no clasificado < confidencial < secreto < muy secreto

Así, las categorías de objetos están formadas considerando combinaciones de características elementales sobre objetos. Por lo general se pueden considerar cuatro grandes categorías:

- Libre
- Restringido
- Sensitivo
- Codificado

El protocolo de control de flujos en este caso se puede presentar de la siguiente forma:

Un flujo de información del nivel de seguridad (i,x) al nivel de seguridad (j,y) es posible si:

1 . $i < j$
2. $x \subseteq y$

Ejemplo. Los siguientes flujos de datos son posibles:

[Secreto,{Restringido,Sensitivo}] → [Muy secreto, {Restringido, Sensitivo}].
[Secreto, {Sensitivo}] → [Secreto, {Restringido, Sensitivo, Codificado}].

Los siguientes flujos de datos no son posibles:

[Secreto, {Restringido,Sensitivo}] → [Libre, {Restringido, Sensitivo}].
[Secreto, {Restringido,Sensitivo}] → [Secreto, {Restringido}].

CONTROL DE LA INFERENCIA

Otra de las técnicas utilizadas para mantener la seguridad de una base de datos es por medio del control de la inferencia, es decir,

evitando que una persona pueda, sin tener autorización, conocer e inferir información, a partir de los datos a cuales sí tiene acceso. Ejemplo. Control sobre el tamaño del resultado:

Pregunta ¿Cuántos enfermos hay con las siguientes características?

- Sexo masculino
- Edad 45-50
- Casado
- Dos niños
- Graduado de Facyt-Computación.
- Gerente de la Empresa X.

Respuesta: Suponer que la persona que hace esta pregunta sabe que es Martínez quien posee estas características. Ahora, desea descubrir informaciones confidenciales sobre Martínez, por ejemplo.

Pregunta: ¿Cuántos enfermos hay con las siguientes características?

- Sexo Masculino
- Edad 45-50
- Casado
- Dos niños
- Graduado de Facyt-Computación.
- Gerente de la Empresa X.
- Toma calmantes debido a depresiones

Respuesta: La respuesta será 1 si Martínez toma calmantes, 0 en caso contrario. Así, la persona no autorizada encuentra una condición C en donde el conjunto resultante solo tiene un elemento.

Por lo tanto, ¿cuántos individuos satisfacen C y X? Si la respuesta es 1 el individuo posee la característica X, 0 en caso contrario. Así, la regla que debe respetarse es la siguiente:

No responder a las consultas en donde hay menos de k o más de n tuplas en la respuesta. En este caso n es el número total de tuplas de la base de datos y k es un entero positivo que especifica el tamaño mínimo permitido de la consulta.

Si el lenguaje de consultas permite el complemento, un tamaño máximo de n-k debe también controlarse, pues de lo contrario pueden hacer preguntas de tipo NOT C. Sin embargo, este tipo de control no es del todo satisfactorio. En efecto, cuando k es aproximadamente igual a n/2, las fugas son posibles usando la técnica del perseguidor. La idea de base es suponer que a un indiscreto se le informó que una persona P está caracterizada por la condición

C = A AND B

Además, se supone que el SGBD puede responder a las preguntas caracterizadas por la condición A y las caracterizadas por la condición A AND NOT B. La condición

T = A AND NOT B

Se llama el perseguidor de P.

Se llama perseguidor debido a que, quien hace las preguntas, puede usarlo para localizar características adicionales de un individuo.

Ejemplo. Sea la contribución de varias personalidades a un cierto partido político y que se representa en la relación. Por otra parte, suponer que se pueden hacer consultas cuyas respuestas involucran sólo sumas y/o cardinalidades de conjuntos.

CONTRIBUCIÓN

Nombre	Sexo	Profesión	Monto
Alonso	m	Periodista	30000
Bernardo	m	Periodista	5000
Carlos	m	Empresario	1000
Sara	f	Periodista	50000
Sonia	f	Profesora	1000
Tomás	m	Profesor	20000
Uriel	m	Médico	200000
Verónica	f	Abogada	100000

Suponer que C = (periodista AND femenino) identifica en forma única a Sara.

Con el perseguidor:
T= (periodista AND NOT femenino) = (periodista AND masculino),
se conocerá la contribución de Sara:

P: ¿Cuántas personas son periodistas?
R: 3
P: ¿Cuántas personas son periodistas y hombres?
R: 2

Así, C identifica a Sara.

P: ¿,Cuál es la contribución total de los periodistas?
R : 85000
P: ¿Cuál es la contribución total de los periodistas, varones?
R: 35000

Por lo tanto, Sara contribuyó con 50000 unidades.

LA CRIPTOGRAFÍA

Consiste en almacenar o transportar la información en forma tal que solo los usuarios que posean el código sean los que puedan comprenderla. Tanto para el caso de los controles de flujo, de acceso o de inferencia no se puede evitar que un operador deje en su "mesa de trabajo" una lista de *passwords* de usuarios o que un indiscreto vaya a consultar las informaciones que pueden estar en algún dispositivo de respaldo.

Para evitar este tipo de indiscreción se puede cifrar la información de tal manera que solo pueda ser descifrada por personas -individuos o programas- que conozcan el código secreto. Este método de cifrar mensajes se conoce como criptografía y es muy seguro para el almacenamiento y transporte de la información.

Principales elementos de un sistema criptográfico

- En primer lugar, se tiene lo que se llama el código que es un método secreto de escritura, en donde el mensaje que se desea almacenar o transportar en una red es codificado y se llama criptograma.
- El proceso de transformar un mensaje en uno codificado se denomina encriptamiento.
- Por otra parte, el proceso inverso de transformar un criptograma en un mensaje se denomina desencriptamiento.

Ejemplo:

De una manera más formal, se puede decir que un sistema criptográfico consta de cinco componentes:

- Un mensaje M que debe almacenarse o transportarse.
- Un criptograma C
- Un conjunto de claves I
- Un conjunto de reglas de encriptamiento
  e_i: :$M \rightarrow C$ con $i \in I$
- Un conjunto de reglas de desencriptamiento

$d_j : C \rightarrow M$ con $j \in I$

Como ejemplo de un algoritmo de encriptamiento, se puede mencionar el algoritmo de cifrados de transposición. Este tipo de algoritmo reordena los caracteres que componen el mensaje según un esquema establecido.

Así, por ejemplo, el algoritmo de transposición columnar, toma un mensaje y lo escribe en una matriz m x n y el criptograma es el que se genera leyendo el mensaje en forma columnar.

Ejemplo. Considerar el mensaje: ser o no ser; he ahí la cuestión y se desea utilizar una matriz de n x 5. En este caso, dicha oración queda ubicada de la siguiente forma:

s	e	r		o
n	o		s	e
e	r	,		h
e		a	h	í
	l	a		c
u	e	s	t	i
ó	n			

Así el criptograma sería el siguiente:

s ee uó enr len ro,aas ht oshíci.

Por supuesto que dicho algoritmo es muy sencillo y es relativamente fácil de descifrar. Sin embargo, existen otros tipos de algoritmos más complejos y que permiten alta confiabilidad. Este es el caso de los métodos de encriptamiento llamados de *clave pública*. En este caso el código de encriptamiento no puede invertirse fácilmente para hallar el código de desencriptamiento.

El sistema de encriptamiento público más conocido y usado en la actualidad es el llamado RSA (R. Rivest, A. Shamir y L. Adlema). La seguridad de este algoritmo radica en el problema de la factorización

de números enteros. Los mensajes enviados se representan mediante números, y el funcionamiento se basa en el producto, conocido, de dos números primos grandes elegidos al azar y mantenidos en secreto. Actualmente estos primos son del orden de 10200, y se prevé que su tamaño aumente con el aumento de la capacidad de cálculo de los ordenadores.

Como en todo sistema de clave pública, cada usuario posee dos claves de cifrado: una pública y otra privada. Cuando se quiere enviar un mensaje, el emisor busca la clave pública del receptor, cifra su mensaje con esa clave, y una vez que el mensaje cifrado llega al receptor, este se ocupa de descifrarlo usando su clave privada.

Se cree que RSA será seguro mientras no se conozcan formas rápidas de descomponer un número grande en producto de primos. La computación cuántica podría proveer de una solución a este problema de factorización.

El método se basa en las dos suposiciones siguientes:

- Es relativamente fácil determinar dos números primos grandes p y q y multiplicarlos para obtener un número n = p * q que sirve de base para un código de encriptamiento.

- El proceso inverso, es decir, determinar los números primos p y q a partir del número n es difícil realizarlo en un tiempo aceptable cuando n es un número muy grande.

Ejemplo: María desea enviar un mensaje a Carlos, quien escoge dos números primos p y q cuya división por 3 tenga a 2 como resto. Así, calcula n = p*q y se lo envía a María.

Por su parte, María transforma el mensaje en una sucesión de dígitos -lo cual resulta muy fácil- y la divide en segmentos de longitud n. Sea uno de esos segmentos representado por el número x. María procede a calcular el resto y al dividir. X^3 por n y el segmento codificado es el número y.

Cuando Carlos recibe el mensaje utiliza un número ultra secreto e = (2(p-1)(q-l) +1)/3 el cual es entero por las escogencia; de los números primos p y q. El fragmento original por su parte, es decir el número x, es igual al resto de dividir y por n. En este caso a una persona le será imposible descifrar en un tiempo oportuno el criptograma, aunque conozca el número n.

Así, aunque se conozca el número n, se requiere factorizar p-q con el fin de descifrar el mensaje. Sin embargo, si los números primos escogidos son muy grandes -de varios cientos de dígitos-, se

requeriría mucho tiempo de máquina -inclusive en los computadores más rápidos para lograr tal factorización.

Actualmente, se considera que la técnica para seguridad de la información más barata y segura es la criptografía.

REGLAS DE INTEGRIDAD

Antes se han estudiado algunas reglas de integridad, como la integridad de relación, integridad de dominio e integridad referencial. Existe otro tipo general de reglas de integridad y que podrían considerarse como generalización de las anteriores y que se denominan (operaciones de disparo, *triggers*).

Las operaciones de disparo son reglas que regulan la validez de insertar, suprimir, actualizar y recuperar operaciones, incluyendo los efectos de éstas sobre otras relaciones u otros atributos dentro de la misma relación.

Ejemplo. Requerimientos que generan operaciones de disparo:

- Un cliente no puede tener un crédito superior a Bs. 4000000.
- Una empresa no puede exceder el límite de crédito establecida para cada uno de sus clientes.
- Las unidades solicitadas en una orden de cliente deben ser múltiplos del número de unidades por paquete del producto.

Las operaciones de disparo reflejan tipos de restricciones que pueden involucrar:

- Varios atributos de diferentes relaciones.
- Dos o más atributos en una misma relación.
- Un atribulo o relación y un parámetro externo.

Las operaciones de disparo tienen dos componentes lógicos. Por una parte, se tiene el disparador, o evento y condición que causa que la operación suceda y la operación, o la acción de disparo. Cuando se establece una operación de disparo es recomendable considerar al menos los siguientes componentes:

- Evento que provoca la operación de disparo: inserción, actualización, supresión o recuperación.
- Objeto del evento: nombre de la relación y/o atributo que va a ser modificado o accedido.
- Condición bajo la cual se inicia la operación de disparo
- Acción que se lleva a cabo -tal como evento rechazado o evento asociado al disparador-

Ejemplo. La operación de disparo generada por el primer requerimiento en el ejemplo anterior puede representarse en un diccionario de datos de la siguiente forma:

Requerimiento	Evento	Relación	Atributo	Condición	Acción
Un cliente no puede tener un crédito superior a Bs. 4000000.	Inserción	CLIENTE	Crédito	Si monto es superior a 4000000	Rechazar la inserción

Referencias

• Date, C. J. 1993. Introducción a los Sistemas de Bases de Datos. Vol. 1 (Ed. Addison-Wesley Iberoamericana).
• Fernandez, E. B., Summers, R. C. y Wood, C. 1981. Database Security and Integrity (Addison-Wesley).