PROTECTION AND SECURITY OF DATA BASE INFORMATION SUMMARY

Mariuţa ŞERBAN, Assist. Lecturer PhD
Spiru Haret University
t_mariuta@yahoo.com

Abstract

Data bases are one of the most important components in every large informatics system which stores and processes data and information.

Because data bases contain all of the valuable information about a company, its clients, its financial activity, they represent one of the key elements in the structure of an organization, which determines imperatives such as confidentiality, integrity and ease of data access.

The current paper discuses the integrity of data bases and it refers to the validity and the coherence of stored data. Usually, integrity is defined in connection with terms of constraint, that are rules regarding coherence which the data base cannot infringe.

Data base that integrity refers to information correctness and assumes to detect, correct and prevent errors that might have an effect on the data comprised by the data bases.

Key-words: *data security, data base security, data protection, integrity restrictions, control*

JEL Clasification: C_{61}, C_{63}, C_{88}, P_{41}

Introduction

Because data bases contain all of the valuable information about a company, its clients, the financial activity, they represent one of the key elements in the structure of an organization, which determines imperatives such as confidentiality, integrity and ease of access to data. These can come in the form of management systems for documents and accounting, billing systems, content management systems (CMS), management systems for technological processes of production.

The risks that can appear in this case come in the form of both external (viruses, hackers, competitors etc.) and internal (theft, mal-intended use, errors and a weak surveillance over personnel and users) factors. Also, it is necessary to take into consideration the reliability of systems, SGBDs, technical means that can interact with some factors like: fires, natural calamities.

Therefore, initially, in order to find solutions, it is necessary to establish the weak points of the data base:

1. The SQL language – it represents a powerful tool of data interrogation, but it also permits the hacking of the system, meaning:

– accessing information using logical deduction, matching or breaking the passwords of the data base users, breaking the system in order to increase the number of privileges;

– data aggregation;
– modifying or replacing data etc.

2. Access management to SGBD/BD:
– weak surveillance over personnel and users;
– errors determined by users' mistakes or those of external users;
– equipment theft;
– information breakage.

3. Attacks against the data base information:
– information tapping while it is circulating through the network;
– the use of the existing network connections to the data base, established by logged in users.

4. Other types of attacks:
– operating system attacks;
– buffer overburdening, blocking data access;
– hackers and viruses.

Data integrity in models of data base information organizing

Data base security can be insured by applying some models that correspond more or less to the type of activity, security policy and to the level of importance of the company's information.

A simple model of data organizing is composed by two elements:

– access control – where every user or informational process of the system has a certain set of permitted actions, which can be executed in relation to certain objects;

– authenticity control – it checks if the user or the process is indeed trying to perform the action it is being entered in the system.

A more complex model of data organizing is the data base multilevel security one, which represents a powerful instrument. This one also has some downsides regarding productivity, costs and ease of access. In these types of systems, the information is organized in several degrees of importance and it usually uses the Bell-LaPadula model that manages subjects, processes and objects.

The methodologies of fighting against data base risks and attacks can be ranged into three categories:

- Data base and SGBD security – scanning the data base, auditing the vulnerabilities, monitoring the activities, control over data base access.
- Auditing network and operating system vulnerabilities – mechanisms that are based on scanning the network and outside. The results consist of examinations, reports and identifying the weak spots.
- Encrypting the data base – a management system of keys that allow a variety of encrypting algorithms.

Data integrity in data base management systems

Data base integrity refers to information faultlessness and it implies identifying, correcting and preventing different errors that can affect the data base information. When referring to data base integrity, this means that the data is

consistently relative to all of previously established restrictions (that apply to that data) which make the data valid. The integrity conditions, known also as integrity rules or restrictions, do not allow users to add to the data base abnormal data or that are reflected through data applied criterions.

There are many criterions when making an integrity rules classification. There are a series of structural rules, connected to certain levelness between values and these are expressed through functional dependency or through setting fields of unique values.

Another series of conditions are structured by the unit where the restriction is applied, and in this case there are domains restrictions (regarding certain values for attributes) or tables restrictions (connections). Table based restrictions can be unituple (it refers to each tuple separately) or multiple (it refers to different tuples combinations).

An integrity restriction of unituple connections imposes that in every tuple of a base connection, in the fields corresponding to the primary key, to have values that are different from the correspondent null values. This rule can also be stated as: “in a relational data base there is no information enlisted regarding something that cannot be identified”.

A multituple integrity restriction example is the referential restriction which implies the condition that the external keys, if they are not null, they will have values corresponding to one of the primary keys existing in the referred connection. This condition is checked every time a new tuple that contains an external key is being inserted or the values of a tuple’s external key are being modified, signaling the possible discrepancies canceling the modifications. Other such examples are to verify the uniqueness of the primary key and the restrictions resulted from the functional and multivalue dependence.

Another classification criterion is the one that decerns between integrity rules that refer to different statuses of the data base from the rules of transitioning from one status to another.

Restrictions can be structured also by taking into consideration the moment in which they are applied, thus having immediate rules (that are verified when the indicated action is being executed) or postponed rules (that are verified after other connected actions have been executed but not before modifying the data base).

According to the applicability area, restrictions can be grouped in general restrictions, which are applicable to all data base connections and specific restrictions, which are applicable to only certain connections.

Integrity rules are applicable to data base connections that have represented the effective information of the data base. Some of the most frequently used general rules in the relational model are the rules regarding primary keys (the uniqueness of the connection’s primary keys) and external keys.

Two types of integrity restrictions can be analyzed:

1. *Restrictions regarding entity integrity*

In a base relation the attribute of a primary key cannot be null. There is known that (in many situations) the values of the primary keys must be unique. For

most SGBD the uniqueness of the primary key and the integrity of the entity are compulsory conditions.

2. *Restrictions regarding referential integrity*

If in a connection there is an external key then either the value of the external key must match the value of a competing key from a certain tuple in the originating connections, either the value of the external key must be null.

In other words, if a value exists in a connection as an external key, either it has to match the value of an external key from another connections either it has to be null.

Strong issues can appear when the values of the primary keys should be modified or deleted.

If a primary key is being updated or if the entire tuple is being deleted and if

– the value of the primary key does not appear anywhere as an external key then the action can be performed;

– the value of the primary key appears elsewhere as a primary key then the action cannot be performed;

– The action can be performed but there are also set up the entries of the external key at null values or at an implicit value (if one was mentioned);

– The action can be performed but also:

– *in the case of updates* – the changed value is propagated to the external key's advents where the external key is a part of the primary key of the connection and also the changes will be conveyed where the mentioned primary key is an external key in another connection.

– *in the case of erasures* – the growth is propagated, meaning that there will be deleted the tuples that have values of the external key matching the primary key and that communicate with the user.

All of the above are general rules that can, according to each case, suffer small transformations related to a specific application.

The situation described above can be solved in the application or can be included in the SGBD using the triggers mechanism. Most SGBD allow the creation and registering of data base procedures, procedures that can be invoked when certain events occur (such as updating or deleting attributes).

Data protection

Data protection against failure events such as system failure caused by software or hardware components breakdown is also associated with data integrity.

Accidental data loss of consistency can be caused by:

– incidents that occur when transactions are being performed;

– anomalies due to concurrent access to the data base;

– anomalies due to working with data distributed on several computers in a network;

– logical errors generated by the programs belonging to the users and others.

For the restructuring of data bases when inconsistencies can occur, most data bases are being periodically copied using magnetic means kept in safe places. There is also kept an inventory of all the transformations performed on the data

base since the last copy has been made. The file that contains these modifications is known as a journal. Each record from the journal contains an identifier of the program that has performed the modification, the former value and the new value of an item. In the journal there are also kept different important moments in the functioning of a program (initiating and ending a program, ending certain operations etc). All of the mechanisms mentioned in this paragraph are the features of working with transactions. At the completion of a transaction, regardless of the fact that it ended normally or with errors, the data base must have the same degree of integrity as the one from the moment when that certain transaction has been initiated.

A transaction has been executed if all of the reckonings generating by it in the working area have been completed and a copy of its results has been made in the journal. The occurrence of failures after the transaction has been executed does not have an impact on the data base because the data base can be rebuilt with the help of the last copy and of the journal which contains all the results of the executed actions. The modifications generated by unfinished or uncommitted transactions are not taking into consideration when rebuilding the data base by using the journal.

The committing strategy contains two phases as it follows: a transaction can write in a data base only after it has been committed and a transaction can be committed only after it has registered in the journal all the item changes generated by it.

There are a few aspects that should be mentioned regarding the integrity and security of data bases:

- Making sure that stored up data coherency related to their own significance. For example, the cornered quantity must not be negative; the paid salary should not be smaller than the minimum guaranteed salary etc. The rules of semantic and structural integrity should be respected.
- Making sure that the concurrent actions of users do not bring prejudices to other users. This refers to the synchronization of concurrent access to the data base.
- Making sure that after a physical error (power failure etc.) the data base will remain coherent. This refers to safety in data base functioning.
- Making sure that the data base is being used by users that have that right. This refers to assuring security means when using the data base.

Usage security

SGBD must have mechanisms that insure the security of the stored information. Below there are mentions four groups of techniques that insure security:

- the control of access – it verifies the identity of the users and their access rights;
- the control of data flows – it monitors the track of the data in order to protect it from being used inappropriately;

- the control of inference – the user cannot construe confidential information through the interface (using the data that the users has access to).

Cryptography has the purpose of stocking and transporting data in a manner understood only by the users that have a certain code to translate it. With the help of a secret code C, the data D is encrypted (using an encrypting program) and can be decrypted only by the holder of that secret code. There are several programs that perform encrypting/decrypting actions (the PGP program that performs on the Windows platform).

The issue regarding the security of usage becomes more and more important as the data bases informational system are more open to the internet.

Assuring the security of data refers to the protection of data bases against their unauthorized usage and especially against unwanted data modifications or obliterations and against unauthorized access to data.

Technical and administrative controls are used for assuring the security of the data base information;

Usually, security is associated with the following situations:

- unauthorized access to data;
- losing data confidentiality (secrecy);
- losing data privacy;
- losing data integrity;
- losing data availability.

It is challenging to protect data from mal-intended access. It is admitted the fact that there is no absolutely safe data protection, but only protection and security measures more or less efficient.

Examples of on purpose mal-intended access to a data base information:

- unauthorized reading of data;
- unauthorized modifications of data;
- data deleting.

The term of data base security is usually associated with the mal-intended access, while integrity refers to avoiding accidental losses of consistency.

For data base protection there can be taken integrity assuring measures at different levels:

- at physical level – the computers should be placed in a space that does not allow the access of unauthorized people;
- at people level – it is recommended to grant access authorizations with lots of care and to keep a strict track of all the authorized users;
- at operating system level – the shortness of the operating system should be eliminated or compensated by other measures;
- at SGBD level – the systems provides certain facilities that assure data protection.

Techniques of data security insuring

Users identification

Each user is being granted certain rights to operate din specific areas of the data base at different levels such as the connection, registration, the page, the attribute etc. These rights refer to the possibility of reading, inserting, deleting or modifying specific data. Identification is usually done through passwords set by either the data base administrator or by the user.

Data protection through coding (encrypting)

Because there are other ways besides through a SGBD to reach data (for instance, through directly reading the magnetic environment) protection can be done by keeping data encrypted on the magnetic environment.

The use of views in applications

The ability of views to 'hide' some of the information from the data base can be also used for setting a certain degree of data security. Therefore it exists access at connection (table) level or access at view level.

Some systems do not accept modification made through views. These views are called read-only and are used especially for applications in which data can be read by all users (public data bases), but the modifications are made only with the administrator/data base owner approval.

The views modifications are usually not permitted because these can lead to lateral effects over parts of the data base that do not appear in views. For instance, deleting an item from a view can lead to the elimination to another item that has a connection with the deleted item and is not included in the view.

Rights administration and transfer

A strict access rights inventory is being kept for each user at certain areas of the data base and there are set rules for access right transfer from one user to another.

Means of authorization:

- authorization for reading (consulting);
- authorization for inserting (adding);
- authorization for updating (excluding errors);
- authorization for deleting (at tuple level).

For the above cases it is not take into consideration the issue of performing changes at data base scheme level. If this aspect is included then we have the following:

- authorization at index level (creating and deleting indexes);
- authorization at connections level (creating);
- authorization for modifications at connections level (deleting or adding attributes in connections);
- authorization for deleting all connections.

Different protection methods can be indicated through the language used for data processing. The data base areas that can be accessed by the user can be

defined through selection and protection actions which make visible other areas of the data base.

The management of the organization, which is the data base owner, must take security measures that will reduce the risk to lose or destroy information. Through information loss it is understood that the private sense of the information is being lost, becoming available for a larger group of persons than the initial one. The "leaks" of information don't always leave traces, therefore they cannot always be found as detectable changes in the data base.

Conclusions

The problematic of security refers to legal, social and ethic aspects, physical control aspects (guarding and the possibilities to block access at terminals), access aspects (conditions for access approval), operational aspects (passwords creation), hard control aspects (the way of hard access to different components), operating system security (information protection and canceling intermediate results in order to keep data secrecy), aspects regarding the notion of data ownership and other similar ones.

There must be mentioned that thefts and frauds are not necessarily related to the data base. They represent a risk factor for the entire organization, the gravity of this actions varying also according to the companies profile.

In the European Community there has been a serious concern in order to bring to date the specific legislation in order to serve the current needs generated by the high use of computers. What is mainly aimed is to establish laws that will protect both persons and organizations and that will respond to the need of information privacy, meaning that the information should not be available for access to a larger or smaller audience if this could have negative consequences on the owner of those pieces of information.

REFERENCES

Băjenescu, T.I. (2003), *Progresele informaticii, criptografiei şi telecomunicaţiilor în secolul 20*, Editura Matrix Rom, Bucureşti.

Burtescu, E. (2002), *Securitatea bazelor de date distribute*, Catedra de Informatică Economică, ASE, referat doctorat.

Fusaru, D. (2002), *Arhitectura bazelor de date. Mediul SQL*, Editura Fundaţiei *România de Mâine*, Bucureşti.

Florescu, V., Năstase, P. (2002), *Baze de date. Fundamente teoretice şi practice*, Editura Infomega, Bucureşti.

Ivan, I., Toma, C. (2006), *Informatics Security Handbook*, Editura ASE, Bucureşti.

Wikipedia, Information Security, http://en.wikipedia.org/wiki/Information_security.

www.betabuzau.ro/uploads/resurse_pdf/7Securitatea.DOC.

Wikipedia, IPSec, http://ro.wikipedia.org/wiki/IPSec.