IRIA

laboria

laboratoire de recherche
en informatique
et automatique

PROTECTION OF INFORMATION IN RELATIONAL DATA BASE MANAGEMENT SYSTEMS

François M. BANCILHON
Nicolas SPYRATOS

Rapport de Recherche N° 243

Août 1977

Institut de Recherche
d'Informatique
et d'Automatique

ine de Voluceau
ocquencourt
8150 - Le Chesnay
France
Tél.: 954 90 20

PROTECTION OF INFORMATION IN RELATIONAL DATA BASE MANAGEMENT SYSTEMS

François Bancilhon*, Nicolas Spyratos**

Résumé :

Nous étudions ici le problème de la protection des informations dans une base de donnée relationnelle. Nous supposerons que l'on se protège d'un usager dont le seul moyen d'accès à la base est un langage de requêtes relationnel.

Le but de cet article est de formaliser la notion de protection. Nous décrivons donc d'abord un système de protection de manière informelle, puis nous présentons un modèle formel pour ce même système. Les entités à protéger sont des propositions qui ont été déclarées confidentielles. Une requête de l'usager «viole» une proposition protégée si la réponse à cette requête modifie la «connaissance» que l'usager a de cette proposition. Pour finir, nous démontrons que ce modèle est un bon outil d'analyse et d'évaluation des systèmes de protection.

Abstract :

This paper is concerned with protection of information in a data base from disclosure to properly identified users. It is assumed that the only means of access the user has is through a relational query language.

The objective of the paper is to formalize the notion of protection. This is done by first describing intuitively a protection system and then by proceeding to present a formal model for this system. The objects to be protected are propositions that have been declared confidential. A user query violates a protected proposition if the answer modifies the knowledge the user has about this proposition. Following this approach, we propose some protection systems and we discuss their implications. It would seem, however, that designing a «perfect» protection system is an almost impossible task.

* IRIA/LABORIA.

** IRIA/LABORIA – Post-Doctoral Fellow of the National Research Council of Canada.

1 - INTRODUCTION

The increasing amount of valuable and private information processed by computers implies a longterm need for rigorous controls.

Personal information, as defined by privacy legislation, and financial or medical records, are examples of information which must be protected during computer processing. It is hard to give a precise definition of computer security because specific security requirements depend so strongly on the larger human, social and financial systems that are served by the computer processing. In general, "security is concerned with any unauthorized or undesirable modification, disclosure or destruction of information" [Linder 76] .

Security must be concerned with any path by which information could be modified, disclosed, or lost. However, some aspects of the overall security problem are clearly beyond the control of a central computer system and can be regarded as separate problems. For example, communication security, identification of users, and physical protection of the computer site are distinct problems. Also, the security of most systems can easily be broken if an operator can be bribed. This might seem to be outside the control of the hardware/software system. However, if a system is designed for security, it is reasonable to expect that the operator's command language provides a protection environment that carefully limits his privileges.

This paper focuses on the problem of protection of information from unauthorized or undesirable disclosure. This problem can be very different according to whether we protect information from a non-programmer user, from a team of expert programmers or from a group of spies using bugging devices. It can also vary widely according to whether we protect a record in a file, a file in a software system or a large complex set of integrated files. Thus, it appears that in order to define a protection problem one must answer the following questions :

(1) Against whom do we protect ?

(2) What do we protect ?

(3) How do we protect ?

(4) What does "protect" mean ?

Note that question (4) addresses the problem of defining the protection function whereas question (3) addresses the problem of realizing that same function.

This paper is not concerned with the general study of protection of information. Rather, it focuses on a specific problem, namely, that of protecting information in a data base from disclosure to a properly identified user. It is assumed that the only means of access to the data base by the user is a relational query language.

The objective of the paper is to formalize the notion of protection. This is done by describing first, intuitively, a protection system and the proceeding to present a formal model for this system. Using this model it is shown that even in an "ideal" environment, designing a "perfect" protection system is an almost impossible task.

2 - DEFINITION OF THE PROBLEM

2.1 - Against whom do we protect ?

The problem considered here is that of protecting information against unauthorized or undesirable disclosure. The situation we are interested in is illustrated in Figure 1. On the one hand, we have a data base (which may be hierarchical, network, relational or any other type). On the other hand, we have a user with the following characteristics

(1) The user has been properly identified, i.e., we know who the user is.