Database Security—Concepts, Approaches, and Challenges

Elisa Bertino, *Fellow*, *IEEE*, and Ravi Sandhu, *Fellow*, *IEEE*

Abstract—As organizations increase their reliance on, possibly distributed, information systems for daily business, they become more vulnerable to security breaches even as they gain productivity and efficiency advantages. Though a number of techniques, such as encryption and electronic signatures, are currently available to protect data when transmitted across sites, a truly comprehensive approach for data protection must also include mechanisms for enforcing access control policies based on data contents, subject qualifications and characteristics, and other relevant contextual information, such as time. It is well understood today that the semantics of data must be taken into account in order to specify effective access control policies. Also, techniques for data integrity and availability specifically tailored to database systems must be adopted. In this respect, over the years the database security community has developed a number of different techniques and approaches to assure data confidentiality, integrity, and availability. However, despite such advances, the database security area faces several new challenges. Factors such as the evolution of security concerns, the "disintermediation" of access to data, new computing paradigms and applications, such as grid-based computing and on-demand business, have introduced both new security requirements and new contexts in which to apply and possibly extend current approaches. In this paper, we first survey the most relevant concepts underlying the notion of database security and summarize the most well-known techniques. We focus on access control systems, on which a large body of research has been devoted, and describe the key access control models, namely, the discretionary and mandatory access control models, and the role-based access control (RBAC) model. We also discuss security for advanced data management systems, and cover topics such as access control for XML. We then discuss current challenges for database security and some preliminary approaches that address some of these challenges.

Index Terms—Data confindentiality, data privacy, relational and object databases, XML.

1 Introduction

As organizations increase their adoption of database systems as the key data management technology for day-to-day operations and decision making, the security of data managed by these systems becomes crucial. Damage and misuse of data affect not only a single user or application, but may have disastrous consequences on the entire organization. The recent rapid proliferation of Web-based applications and information systems have further increased the risk exposure of databases and, thus, data protection is today more crucial than ever. It is also important to appreciate that data needs to be protected not only from external threats, but also from insider threats.

Security breaches are typically categorized as *unauthorized data observation*, *incorrect data modification*, and *data unavailability*. Unauthorized data observation results in the disclosure of information to users not entitled to gain access to such information. All organizations, ranging from commercial organizations to social organizations, in a variety of domains such as healthcare and homeland protection, may suffer heavy losses from both financial and human points of view as a consequence of unauthorized data observation. Incorrect modifications of data, either intentional or unintentional, result in an incorrect database state. Any use of incorrect data may result in heavy losses for the organization. When data is unavailable, information crucial for the proper functioning of the organization is not readily available when needed.

Thus, a complete solution to data security must meet the following three requirements: 1) *secrecy* or *confidentiality* refers to the protection of data against unauthorized disclosure, 2) *integrity* refers to the prevention of unauthorized and improper data modification, and 3) *availability* refers to the prevention and recovery from hardware and software errors and from malicious data access denials making the database system unavailable. These three requirements arise in practically all application environments. Consider a database that stores payroll information. It is important that salaries of individual employees not be released to unauthorized users, that salaries be modified only by the users that are properly authorized, and that paychecks be printed on time at the end of the pay period. Similarly, consider the Web site of an airline company. Here, it is important that customer reservations only be available to the customers they refer to, that reservations of a customer not be arbitrarily modified, and that information on flights and reservations always be available. In addition to these requirements, *privacy requirements* are of high relevance today. Though the term privacy is often used as a synonym for confidentiality, the two requirements are quite different. Techniques for information confidentiality

- *E. Bertino is with the Computer Science and Electric and Computer Engineering Department and CERIAS, Purdue University, West Lafayette, IN 47907. E-mail: bertino@cerias.purdue.edu.*
- *R. Sandhu is with the Information Science Engineering Department, George Mason University, Fairfax, VA 22030. E-mail: sandhu@ise.gmu.edu.*

Manuscript received 2 Sept. 2004; revised 11 Jan. 2005; accepted 1 Mar. 2005; published online 4 Apr. 2005.

For information on obtaining reprints of this article, please send e-mail to: tdsc@computer.org, and reference IEEECS Log Number TDSC-0130-0904.