

Preface

The increasing availability of large collections of personal information as well as of data storage facilities for supporting data-intensive services, support the view that service providers will be more and more requested to be responsible for the storage and the efficient and reliable dissemination of information, thus realizing a “data outsourcing” architecture. Within a data outsourcing architecture data are stored together with application front-ends at the sites of an external server who takes full charges of their management. While publishing data on external servers may increase service availability, reducing data owners’ burden of managing data, data outsourcing introduces new privacy and security concerns since the server storing the data may be *honest-but-curious*. A honest-but-curious server honestly manages the data but may not be trusted by the data owner to read their content. To ensure adequate privacy protection, a traditional solution consists in encrypting the outsourced data, thus preventing outside attacks as well as infiltration from the server itself. Such traditional solutions have however the disadvantage of reducing query execution efficiency and of preventing selective information release. This introduces then the need to develop new models and methods for the definition and enforcement of access control and privacy restrictions on outsourced data while ensuring an efficient query execution.

In this book, we present a comprehensive approach for protecting sensitive information when it is stored on systems that are not under the data owner’s control. There are mainly three security requirements that need to be considered when designing a system for ensuring confidentiality of data stored and managed by a honest-but-curious server. The first requirement is *access control enforcement* to limit the ability of authorized users to access system’s resources. In traditional contexts, a trusted module of the data management system is in charge of enforcing the access control policy. In the considered scenario, the service provider is not trusted for enforcing the access control policy and the data owner is not willing to mediate access requests to filter query results. We therefore propose a new access control system, based on selective encryption, that does not require the presence of a trusted module in the system for the enforcement of the policy. The second requirement is *privacy protection* to limit the visibility of stored/published data to non authorized

users while minimizing the adoption of encryption. Data collections often contain personally identifiable information that needs to be protected both at storage and when disseminated to other parties. As an example, medical data cannot be stored or published along with the identity of the patients they refer to. To guarantee privacy protection and to limit the use of encryption, in this book we first propose a solution for modeling in a simple while powerful way privacy requirements through confidentiality constraints, which are defined as sets of data whose joint visibility must be prevented. We then propose a mechanism for the enforcement of confidentiality constraints based on the combined use of fragmentation and encryption techniques: associations broken by fragmentation will be visible only to those users who are authorized to know the associations themselves. The third requirement is *safe data integration* to limit the ability of authorized users to exchange data for distributed query evaluation. As a matter of fact, often different sources storing the personal information of users need to collaborate to achieve a common goal. However, such data integration and sharing may be subject to confidentiality constraints, since different parties may be allowed to access different portions of the data. We therefore propose both a model for conveniently representing data exchange constraints and a mechanism for their enforcement during the distributed query evaluation process.

In this book, we address all these three security requirements by defining a model and a mechanism for enforcing access control on outsourced data; by introducing a fragmentation and encryption approach for enforcing privacy constraints; and by designing a technique for regulating data flows among different parties. The main contributions can be summarized as follows.

- With respect to the access control enforcement on outsourced data, the original results are: the combined use of selective encryption and key derivation strategies for access control enforcement; the introduction of a notion of minimality of an encryption policy to correctly enforce an access control policy without reducing the efficiency in key derivation; the development of a heuristic approach for computing a minimal encryption policy in polynomial time; the introduction of a two-layer encryption model for the management of policy updates.
- With respect to the definition of a model for enforcing privacy protection, the original results are: the definition of confidentiality constraints as a simple while complete method for modeling privacy requirements; the introduction of the notion of minimal fragmentation that captures the property of a fragmentation to satisfy the confidentiality constraints while minimizing the number of fragments; the development of an efficient approach for computing a minimal fragmentation, which is a NP-hard problem; the introduction of three notions of local optimality, based on the structure of the fragments composing the solution, on the affinity of the attributes in the fragments, and on a query evaluation cost model, respectively; the proposal of three different approaches for computing fragmentations satisfying the three definitions of optimality.
- With respect to the design of a safe data integration mechanism, the original results are: the definition of permissions as a simple while complete method for modeling data exchange limitations; the modeling of both permissions and queries as relation profiles and their representation through a graph-based model;

the introduction of an approach for the composition of permissions working in polynomial time; the definition of a method that takes data exchange restrictions into account while designing a query execution plan.

Sara Foresti



<http://www.springer.com/978-1-4419-7658-1>

Preserving Privacy in Data Outsourcing

Foresti, S.

2011, XV, 180 p., Hardcover

ISBN: 978-1-4419-7658-1