Chapter 4
Search in Encrypted Data: Theoretical Models and Practical Applications

Qiang Tang
University of Luxembourg, Luxembourg

ABSTRACT

Recently, the concept of Search in Encrypted Data (SED) has become a highlight in cryptography. A SED scheme enables a client to have third-party server(s) perform certain search functionalities on the encrypted data. In this chapter, the authors conduct a systematic study on SED schemes. First, they describe three application scenarios and identify the desirable security requirements. Second, they provide two orthogonal categorizations and review the related security models for each category of SED schemes. Third, the authors analyze the practical issues related to SED schemes and identify some future research directions.

INTRODUCTION

A Search in Encrypted Data (SED) scheme allows third-party server(s) to search on behalf of a client without the need to recover the plaintext data while preventing the server(s) from learning any plaintext information. SED has become a very active research area in cryptography in recent years. Two seminal SED schemes are the one by Song, Wagner, and Perrig (2000) and the one by Boneh, Crescenzo, Ostrovsky, and Persiano (2004). The first scheme allows a client to encrypt its database and store the encrypted database at a remote server. Later on, the client can instruct the server to search in the encrypted database and return the relevant data. The second scheme is often referred to as PEKS, namely public key encryption with keyword search. With a PEKS

DOI: 10.4018/978-1-4666-4030-6.ch004

scheme, a client publishes his public key so that any entity can encrypt messages for him. Later on, the client can allow a third-party server to search in the encrypted messages by assigning a token to it. Following these two schemes, a lot of variants have been proposed to extend the concepts in many aspects. For instance, Yang, Tan, Huang, and Wong (2010) proposed the concept of public key encryption supporting equality test. In contrast to the scheme by Boneh et al. (2004), the scheme by Yang et al. (2010) allows a third-party server to search on the ciphertexts which are encrypted with public keys from multiple different clients. With the wide adoption of cloud computing applications, SED schemes have been regarded by many to be an important technology in securing outsourcing databases while preserving data utility and confidentiality.

In this book chapter, we aim at a systematic study on existing SED schemes and their security implications. In particular, we reflect on the related theoretical security models and try to understand their practical security guarantees.

In the first step, we study three representative application scenarios, which have motivated a variety of theoretical SED schemes. Despite the frequent citations, the security requirements of these scenarios have not been investigated in depth in the literature. This fact means that there may be a gap between the theoretical security guarantees of existing SED schemes and the practical needs of the application scenarios. In the second step, we present two categorizations for SED schemes. One is based on whether a scheme supports full-domain or index-based search. The other is based on the answers to two questions, namely "Who can contribute searchable data in the outsourced database?" and "Who can issue search queries to the third-party server(s)?" Due to the desired storage and search functionalities, outsourcing data storage and search operations to third-party server(s) inevitably results in some privacy loss for the client. The answers to the above two questions define the characteristics of the search functionalities provided by a SED scheme, and consequently determine the inevitable information leakage of the scheme. In the third step, based on the results in the first two steps, we provide some practicality analysis against the existing provably secure SED schemes. The analysis shows that many practical security concerns have not been covered by theoretical security models. As a result, we are able to identify some future research directions for SED schemes.

Organization. The rest of this chapter is organized as follows. In the second section, we describe three application scenarios and identify their security requirements. In the third section, we categorize the existing SED schemes. In the fourth section, we review the security models for SED schemes. In the fifth section, we analyze the practical issues facing SED schemes. In the sixth section, we conclude the chapter.

SED APPLICATION SCENARIOS

In this section, we describe three representative application scenarios for SED schemes and the related security requirements. In addition, we also mention some possible variants.

Search in Outsourced Personal Database

Suppose Alice is a frequent traveler and needs to access her database during her travel anytime and anywhere in the world. For this, Alice can outsource her personal database to a third-party service provider, such as Google or Dropbox. With this approach, Alice needs to reveal everything (the data and search criteria) to the third-party service provider, which makes the solution undesirable from the privacy perspective. To achieve a privacy-preserving solution, Alice can employ a SED scheme to encrypt her database and outsource the ciphertext. Later on, Alice can issue a search query (containing encrypted search criteria) to the service