Chapter XII
Privacy Preserving Data Mining, Concepts, Techniques, and Evaluation Methodologies

Igor Nai Fovino
Joint Research Centre, Italy

ABSTRACT

Intense work in the area of data mining technology and in its applications to several domains has resulted into the development of a large variety of techniques and tools able to automatically and intelligently transform large amounts of data in knowledge relevant to users. However, as with other kinds of useful technologies, the knowledge discovery process can be misused. It can be used, for example, by malicious subjects in order to reconstruct sensitive information for which they do not have an explicit access authorization. This type of "attack" cannot easily be detected, because, usually, the data used to guess the protected information, is freely accessible. For this reason, many research efforts have been recently devoted to addressing the problem of privacy preserving in data mining. The mission of this chapter is therefore to introduce the reader to this new research field and to provide the proper instruments (in term of concepts, techniques, and examples) in order to allow a critical comprehension of the advantages, the limitations, and the open issues of the privacy preserving data mining techniques.

INTRODUCTION

We live today in the information society. Every second, millions of information are stored in some "Information Repository" located everywhere in the world. Every second, millions of information are retrieved, shared, and analyzed by someone. On the basis of the information stored in a database, people develop economical strategies and make decisions having an important effect on the lives of other people. Moreover, this information is used in critical applications, in order to man-

age and to maintain, for example, nuclear plants, defense sites, energy and water grids, and so on. Information is a precious asset for the life of our society.

In such a scenario, information protection assumes a prominent role. A relevant amount of information stored in a database is related to personal data or, more in general, to information accessible only by a restricted number of users (we call this information "Sensitive Information"). Let us consider as an example the case of a Hospital Health Database. In such a database, records are collected related to the patients of the hospital. The data stored in such database are extremely useful; in fact they allow keeping track of the medical history of the patients, to make an automatic profile analysis, to extract statistical data related to a certain disease, and so on. However, such data can even be considered extremely sensitive. For example, the information "*Patient A has been, in the past, affected by the psychological problem Y*" is an information which, if freely accessible, could have a strong impact on the social life of Mr. A.

It is evident that the concept of *Information Privacy* is a not negligible issue in this context. In the scientific literature, several definitions exist for privacy. At this moment, in order to introduce the context, we briefly define privacy as *limited access to a person and to all the features related to the person*. In the database context, the privacy property is usually satisfied by the use of access control techniques. This approach guarantees a high level of privacy protection against attacks, having as the final goal the direct access to the information stored in a database. Access control methods, however, result nowadays prone to a more sophisticated family of privacy attacks based on the use of data mining techniques. Data mining (DM) techniques has been defined as "The nontrivial extraction of implicit, previously unknown, and potentially useful information from data" (Frawley, Piatetsky-Shapiro, & Matheus, 2002). In other words, by using DM techniques, it is possible to extract new and implicit information from known information. This characteristic constitutes, per se, an enormous advantage in the analysis of immense datasets. However, the malicious use of such techniques is a serious threat against privacy protection.

In a typical database, a large number of relationships (both explicit and implicit) exist between the different information. These relationships constitute a potential privacy breach. In fact, by applying some access control methods, one can avoid the direct access to sensitive information. However, sensitive information, by the presence of these relationships, influences in some way or other information. It is then possible, by applying DM techniques to the accessible information to reconstruct indirectly the sensitive information, violating in such a way the privacy property.

Let us consider the previous Health Database example. In such a case, we can make the hypothesis that only authorized personnel have full access to all the data stored in the database. However, considering that such data can be useful even for some analysis based on statistics, we can imagine a control access policy which allows different levels of access; that is, there exist different user profiles which can access different portions of data. Such a scenario is very common in the real world and guaranteed to avoid the *direct* access to a target data by non-authorized people. However, as claimed previously, due to the relationships among the different data contained in a database, one, by the use of data mining techniques, someone may be able to indirectly infer sensible data starting from the analysis of the public data.

Recently, a new class of data mining methods, known as privacy preserving data mining (PPDM) algorithms, has been developed by the research community working on security and knowledge discovery. The aim of these algorithms is the extraction of relevant knowledge from large amounts of data, while protecting at the same time sensitive information. The main scope of this chapter is then to give a high level overview