Privacy Preservation in Information Systems

Debanjan Sadhya
IIT-BHU, India

Shekhar Verma
IIIT-Allahabad, India

INTRODUCTION

With the dawning of the modern scientific world, technology is gradually becoming more and more complex. Along with the enhancement of technology, the data that is being utilized in it is becoming very much exposed to threats and attacks from potential adversaries and harmful entities. More specifically, the data concerning person specific information is under severe threat because a simple lapse in the security mechanisms in that type of system may lead to very serious problems for the individual, even identity theft. The main responsibility for handling these types of issues lies with the data holders who collect and organise data from individuals for research like purposes. It is upto them to release these information in such a way that privacy or confidentiality is not compromised. Another major issue regarding the privacy controlling mechanism is the utility factor. As discussed later, all major privacy preservation techniques basically modify the original data present in the databases under certain constraints. In this case, care must be taken that the modifications must be made not to diminish the utility or usefulness of the data. Thus a fine balance in data perturbation methods must be struck so that both privacy preservation and usefulness of the data gets preserved in the resulting databases.

To give an concrete example for the need of privacy, let's assume that individuals participate to provide information about their 'name', 'age', 'date of birth', 'marital status', 'past medical history' etc. After gathering all the information, the census micro-data files are made public for research and other related purposes. Now let's assume that the attribute 'past medical history' is a sensitive attribute, i.e. if any value of this attribute is revealed then some harm may happen to the individual . Thus, under no circumstances the value of the sensitive attribute should get revealed because it is fore-mostly important to maintain privacy for the participants and protect their private and personal data from being disclosed. From the point of view of an imposter, he can gradually gain knowledge about his target by firing a systematic sequence of queries on the micro-data databases. The responses of the multiple queries combine to converge towards the sensitive attribute value which was to be concealed. After a number of queries, the final result gets ultimately revealed.

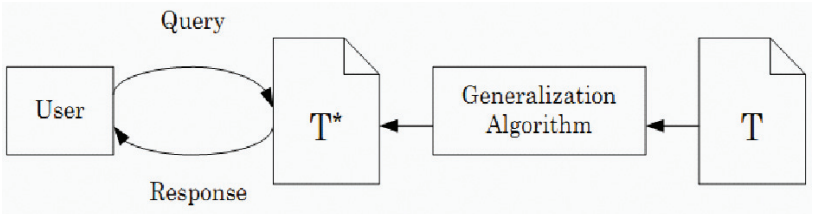
The main objective of the article is to justify the need of privacy preservation techniques and provide a comprehensive overview of the basic methods used for providing the required privacy in information systems. We have already stated the intuition behind these methods (i.e. data perturbation) in this section. In the next section we will discuss in details these procedures and see how privacy is being preserved by their application.

BACKGROUND

The first step in providing privacy to published micro-data is by data sanitization i.e. removing critical attributes related to an individual from the micro-data tables. These attributes are called *identifiers* as they are able to uniquely identify an individual from the database. Some common instances of these types of attributes include social security number and passport number. But in-spite of removing these identifiers, a great number of threats still persist. As shown by Sweeney (2002), even after sanitizing, an individual can be identified by 'linking attack'. In this type of attack, an individual can be identified by correlating

DOI: 10.4018/978-1-4666-5888-2.ch431

Figure 1. Generalization method



shared data amongst multiple databases in which the individual has participated. The attributes which link the individual between the databases are termed as *quasi identifiers.* The medical records of the governor of Massachusetts were easily re-identified by using this linking attack. This fact undoubtedly establishes the high level of danger that persists due to this type of attacks.

To summarize the ideas, the main objective of modern data privacy techniques is to suppress the disclosure risk of individual information as much as possible while maximizing the utility of the presented data. As mentioned by Kiyomoto (2004), there are two main approaches for evading the leakage of personal information from the released micro-data files. Although their methods for achieving privacy are different, but there principal concept is the same i.e. change or modify the original data that is to be released. These two techniques are termed as 'generalization' and 'perturbation'. In the forthcoming sections we will limit our discussions on these broad categories and the techniques implementing them. A more detailed and in depth discussion of privacy preservation techniques is compiled by Fung et al. (2010).

Generalization

Generalization techniques change the original data so that they cannot be identified later on. These methods compute a universal value for a group of records and then replace the individual records within the group with the computed universal value for the group. However, the main drawback for this system is the loss of detailed information incurred during the process. This technique can be visually summarized in Figure 1.

As shown in Figure 1, the original database (T) was transformed to a modified database T* by the generalization algorithm G. Thus the main theme of this type of algorithms is suppression of the original values by substituting with a less detailed but semantically steady modified value. Also, it should also be noted that generalization is a non-interactive technique and is independent of the number and types of queries generated by the user. The user has no complete knowledge of the generalization algorithm as it is partially hidden from him. The main techniques which implement this system include k-anonymity, t-closeness and l-diversity.

Perturbation

Perturbation technique offers an alternate solution for achieving data privacy. The main principle of this system is adding external noise to the original data to produce perturbed output. The nature of the noise added to the original data is generally either Gaussian or Laplacian for reasons discussed later. In spite of retaining the detailed information by the perturbed data, the perturbation technique consists of additional fake information. The complete perturbation method is diagrammatically shown in Figure 2.

As depicted in Figure 2, perturbation methods are basically interactive methods. The answers to the queries fired by the users are calculated in a runtime environment (instead of storing a predefined database as in the case of generalization methods). The original database is represented by 'T', the perturbation algorithm by 'K', the random value by 'rand' and the user generated queries by 'query'. Basically, the original data-set values are dynamically modified by a randomly generated noise which in turn is calculated on the basis of the queries fired by the users. The end results or response can be represented as a function of the perturbation algorithm. Thus, response $(R) = K(T, rand, query)$. The values of the responses to the queries are probabilistic owing to the random nature of the added noise.

Perturbation techniques are broadly classified into two sub-parts namely i) Input perturbation and