Chapter 61

Analytical Study on Privacy Attack Models in Privacy Preserving Data Publishing

Sowmyarani C. N.
R.V. College of Engineering, India

Dayananda P.
JSS Academy of Technical Education, Bengaluru, India

ABSTRACT

Privacy attack on individual records has great concern in privacy preserving data publishing. When an intruder who is interested to know the private information of particular person of his interest, will acquire background knowledge about the person. This background knowledge may be gained though publicly available information such as Voter's id or through social networks. Combining this background information with published data; intruder may get the private information causing a privacy attack of that person. There are many privacy attack models. Most popular attack models are discussed in this chapter. The study of these attack models plays a significant role towards the invention of robust Privacy preserving models.

INTRODUCTION

Data publishing includes 3 entities such as,

1. **Publisher:** The one who publishes the collected data on web.
2. **Data Owner:** The one who owns the data or the data is about that individual.
3. **Data Recipient:** The one who can access the published data.

Publisher will collect the data from data owners. Data owners will have trust over the publisher and give their data. Publisher will publish it by removing the data which may directly leads to breach of

DOI: 10.4018/978-1-5225-7113-1.ch061

privacy of data owners. The data recipients will receive the data from publisher and use the data for analysis purpose to process the data to come out with analogy or decision making.

Publisher before publishing the data should remove the attributes which directly identifies the individual. For example, consider hospital data related to patient entity. Name and complete address will directly identify the individual. So, such attributes will be removed from the data base and rest of attributes will be published. But, this is not sufficient to provide privacy. The person who is interested to know the private information of other individual, can threat privacy of that person. This can be achieved by having background knowledge (Martin, D.J. 2007) about that person and linking the same with the published data. Consider the following data in table form:

Intruder, who is interested in private information of other individual, will have background knowledge (RASTOGI V, 2007) about the person like, where he lives. The area zip code will be 146254. He knows that, that person's age is in 30's. If this background information is linked to the Table 1, intruder can easily conclude that, that person is having hepatitis disease causing a privacy breach. Many privacy preserving methods (Hua-jin Wang 2007, Qinghai Liu 2014) came into existence to avoid privacy breach. Some popular traditional methods are: Perturbation (DUNCAN G 1998, Xiao-Bai Li 2006) and Inference control (Haibing Lu; 2008, Yingjiu Li 2006, R. Brand 2002).

The data in detail which is subject-specific will contain sensitive data which will identify the individuals uniquely. This may lead to violation of individual privacy. There are many laws being enforced to protect privacy. This is the motivation for researchers to work on many privacy preserving models and come up with new techniques which will preserve privacy.

Figure 1. Privacy preserving data publishing

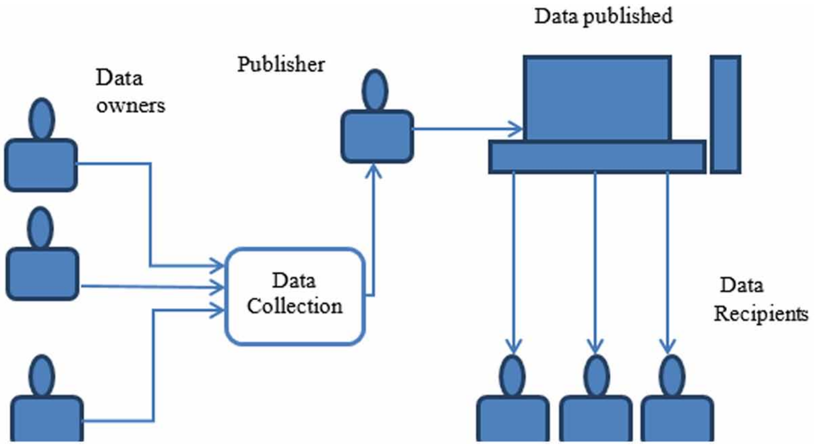

Table 1. Tabular data

Age	Zipcode	Disease
25	146234	Jaundice
35	146245	Gastritis
35	146254	Hepatitis
45	146267	Asthma