INVITED KEYNOTE ADDRESS

Security of Statistical Databases: Overview and Future Directions

Mirka Miller

School of ITMS
University of Ballarat
Australia

m.miller@ballarat.edu.au

Abstract

A statistical database is a database in which the only queries allowed are of statistical type, and based on aggregate data. In particular, a statistical database is not supposed to give answers to queries pertaining to single individual records.

Securing such a database is a difficult problem, since it is often possible to use a clever combination of aggregate queries to derive information about a single individual. If such an inference is possible we say that the database has been compromised.

The security problem for a statistical database is to find suitable control mechanisms so that while statistical information is provided, no sequence of queries is sufficient to infer the values of protected fields of individual records.

Various types of compromise have been defined, including positive, negative and relative compromise, and many types of control mechanisms have been proposed for the protection against database compromise. These mechanisms can be divided into two categories: noise addition and query restriction. However, to date no single security-control method is capable of preventing compromise.

In this talk we sketch the history of the problem, give an overview of the control mechanisms that have been proposed so far, and finally, we consider possible future directions for handling the problem.