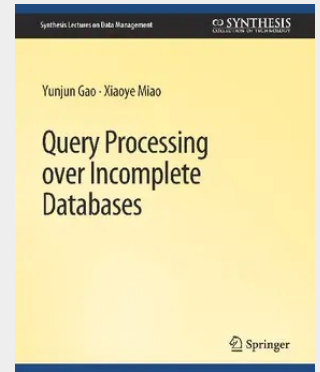


Query Processing over Incomplete Databases

Incomplete data is part of life and almost all areas of scientific studies. Users tend to skip certain fields when they fill out online forms; participants choose to ignore sensitive questions on surveys; sensors fail, resulting in the loss of certain readings; publicly viewable satellite map services have missing data in many mobile applications; and in privacy-preserving applications, the data is incomplete deliberately in order to preserve the sensitivity of some attribute values.

Query processing is a fundamental problem in computer science, and is useful in a variety of applications. In this book, we mostly focus on the query processing over incomplete databases, which involves finding a set of qualified objects from a specified incomplete dataset in order to support a wide spectrum of real-life applications. We first elaborate the three general kinds of methods of handling incomplete data, including (i) discarding the data with missing values, (ii) imputation for the missing values, and (iii) just depending on the observed data values. For the third method type, we introduce the semantics of k -nearest neighbor (k NN) search, skyline query, and top- k dominating query on incomplete data, respectively. In terms of the three representative queries over incomplete data, we investigate some advanced techniques to process incomplete data queries, including indexing, pruning as well as crowdsourcing techniques.

Incomplete data is part of life and almost all areas of scientific studies. Users tend to skip certain fields when they fill out online forms; participants choose to ignore sensitive questions on surveys; sensors fail, resulting in the loss of certain readings; publicly viewable satellite map services have missing data in many mobile applications; and in privacy-preserving applications, the data is incomplete deliberately in order to preserve the sensitivity of some attribute values. Query processing is a fundamental problem in computer science, and is useful in a variety of applications. In this book, we mostly focus on the query processing over incomplete databases, which involves finding a set of qualified objects from a specified incomplete dataset in order to support a wide spectrum of real-life applications. We first elaborate the three general kinds of methods of handling incomplete data, including (i) discarding the data with missing values, (ii) imputation for the missing values, and (iii) just depending on the observed data values. For the third method type, we introduce the semantics of k -nearest neighbor (k NN) search, skyline query, and top- k dominating query on incomplete data, respectively. In terms of the three representative queries over incomplete data, we investigate some advanced techniques to process incomplete data queries, including indexing, pruning as well as crowdsourcing techniques.



58,84 €

54,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9783031007354

Medium: Buch

ISBN: 978-3-031-00735-4

Verlag: Springer

Erscheinungstermin: 20.08.2018

Sprache(n): Englisch

Auflage: Erscheinungsjahr 2018

Serie: Synthesis Lectures on Data Management

Produktform: Kartoniert

Gewicht: 248 g

Seiten: 106

Format (B x H): 191 x 235 mm

