Fuzzy Databases
Modeling, Design and Implementation

JOSE GALINDO
ANGELICA URRUTIA
MARIO PIATTINI

TP311.13
G158

Fuzzy Databases: Modeling, Design and Implementation

José Galindo
University of Málaga, Spain

Angélica Urrutia
Catholic University of Maule, Chile

Mario Piattini
University of Castilla-La Mancha, Spain

IDEA GROUP PUBLISHING
Hershey • London • Melbourne • Singapore

Acquisitions Editor:	Renée Davies
Development Editor:	Kristin Roth
Senior Managing Editor:	Amanda Appicello
Managing Editor:	Jennifer Neidig
Copy Editor:	Eva Brennan
Typesetter:	Amanda Kirlin
Cover Design:	Lisa Tosheff
Printed at:	Yurchak Printing Inc.

Published in the United States of America by
Idea Group Publishing (an imprint of Idea Group Inc.)
701 E. Chocolate Avenue, Suite 200
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@idea-group.com
Web site: http://www.idea-group.com

and in the United Kingdom by
Idea Group Publishing (an imprint of Idea Group Inc.)
3 Henrietta Street
Covent Garden
London WC2E 8LU
Tel: 44 20 7240 0856
Fax: 44 20 7379 0609
Web site: http://www.eurospanonline.com

Copyright © 2006 by Idea Group Inc. All rights reserved. No part of this book may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher.

Product or company names used in this book are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI of the trademark or registered trademark.

Fuzzy databases : modeling, design and implementation / Jose Galindo, Angelica Urrutia and Mario Piattini, editors.
p. cm.
Summary: "This book includes an introduction to fuzzy logic, fuzzy databases and an overview of the state of the art in fuzzy modeling in databases"--Provided by publisher.
Includes bibliographical references and index.
ISBN 1-59140-324-3 (h/c) -- ISBN 1-59140-325-1 (pbk.) -- ISBN 1-59140-326-X (ebook)
1. Database management. 2. Fuzzy sets. I. Galindo, Jose, 1970- II. Urrutia, Angelica. III. Piattini, Mario, 1966-
QA76.9.D3F8935 2005
005.74--dc22
2005013546
British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. Each chapter is assigned to at least 2-3 expert reviewers and is subject to a blind, peer review by these reviewers. The views expressed in this book are those of the authors, but not necessarily of the publisher.

Dedication

To my family and to everyone who finds this book interesting (with a fuzzy degree of at least 0.2).

José Galindo

To my parents, Hugo and Angélica.

Angélica Urrutia

To the Velthuis family.

Mario Piattini

Foreword

Since Zadeh's Fuzzy Sets Theory was formulated, a lot of efforts have been devoted to extend databases with mechanisms to represent and handle information in a flexible way. The proposals appearing in the literature to deal with this aim are mainly supported in the possibilistic models, similarity relationship models, or the combination of both perspectives. This fact, together with the variety of database models susceptible of extension (i.e., relational model, object oriented models, logic model, object-relational model, etc.), has given rise to many approaches of fuzzy database models.

The materialization of these models in Fuzzy DBMS has not been so fructuous, and the development of applications supported by these systems is in an exploratory stage.

The implementation of Fuzzy DBMS will be determined by the development of applications that take advantage of the capabilities of these ones to operate with flexible information when solving real-life problems. In this sense, different areas of application have appeared, and in this book, some examples are collected, such as data mining, information retrieval, content-based image retrieval, and classical applications in the management field, improved with the possibility of manipulating flexible information (see, for example, http://idbis.ugr.es/immosoftweb for an online real-estate portal based on flexible search. It is built on the FSQL server developed by José Galindo and other members of the IDBIS group).

One issue that, from my point of view, has not been paid enough attention from the scientific community has been the extension of the conceptual models for the design of databases to the ambit of the representation of incomplete

此为试读,需要完整PDF请访问: www.ertongbook.com

information. In this sense, this book put together the most important proposals present in the literature. This study is completed with a deep analysis of the features of modeling susceptible of fuzzy treatment to present, next, a fuzzy extension of the EER model, which gives a notation for each of these features. The fuzzy concepts identified in the ambit of modeling require, in a similar way as in the classical case, a DBMS that permits the representation and handling of this type of information. The authors have incorporated these new characteristics to previous models and prototypes of fuzzy databases. The new model, the new data structures, and the new capabilities of handling have given as a result FIRST-2 and a new extension of FSQL (Fuzzy SQL), both of them thoroughly described in this book. The creation of an algorithm that permits the translation of the conceptual definition in terms of FuzzyEER into FSQL sentences completes an important cycle in relation to the conceptual design oriented to fuzzy databases.

Though the central argument of the book is the description of a notation for the conceptual design in an imprecise environment, this volume collects and proposes many worthy resources in the area of fuzzy databases, which makes it an important reference for those people interested in this field in general.

Dr. Juan Miguel Medina
Senior Researcher
Member of the IDBIS Group
Granada, Spain, January 2005

Preface

In 1965 at the University of California, Berkeley, also called the "Athens of the Pacific," Lotfi A. Zadeh[1] introduced the theory of fuzzy sets and fuzzy logic, two concepts that laid the foundation of possibility theory in 1977. These terms were coined by him to deal with the phenomenon of vagueness, in the cognition process of the human being. According to Zadeh, "the theory of fuzzy sets is a step toward a rapprochement between the precision of classical mathematics and the pervasive imprecision of the real world... a rapprochement born of the incessant human quest for a better understanding of mental processes and cognition[2]."

Since then, an enormous quantity of congresses and publications around the world has intended to explore and develop this basic idea of vagueness and its industrial application. Zadeh also said: "at present, we are unable to design machines that can compete with humans in the performance of such tasks as recognition of speech, translation of languages, comprehension of meaning, abstraction and generalization, decision-making under uncertainty and, above all, summarization of information."

When we look at the growth of the Japanese industry in the 1980s we can understand the relevant impact of "fuzzy technologies" in the modeling and design of new products[3].

In these aspects, the gap between the industrial domain and the research domain can be seen in books, journals, articles, cases studies, proceedings, and so forth. In fact, these are the greatest tools to put the theoretical knowledge in action (i.e., in Idea Group Publishing you can find the latest advance in the research of information science, technology, and management). But it is diffi-

cult to find a pedagogical book to help the learning process of the students in computer science in the area of fuzzy databases.

I am glad to tell you that the book you have in yours hands has the courage to attack the problem of fuzzy databases, with a clear and direct approach guiding the reader step-by-step through the understanding process. Indeed this book has the ability to help you in the modeling, design, and implementation processes of fuzzy databases. This book gives you a first glance at a systematic exposition of the three issues (modeling, design, and implementation). Perhaps the only regret I have in this book is the use of Oracle platform, which, in my view, has the influence of the industrial software of the 1990s. However, the definitions, ideas, and new approaches are platform independent.

Before I say something about the features of the book, I would like to explain some historical aspects that I find interesting to being taken into account by readers. First, in Europe there are two cities well known by the implication of the database in the Zadeh legacy: Toulouse and Granada.

In 1985 Didier Dubois[4] and Henry Prade[5] published *Théorie des Possibilités — Applications à la représentation des connaissances en informatique*, which was translated into English three years later as *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. In Chapter VI of this book the authors introduce the use of the possibility distribution to represent incomplete and uncertain dates in a relational database. This chapter was the result of a PhD thesis written in Toulouse by Claude Testemale[6] and co-directed by Prade. In this work you can see the original code in MACLISP for fuzzy query processing.

Some years later, in Granada, the book of Dubois and Prade, in particular Chapter VI, had a great impact on the PhD thesis of Juan Miguel Medina[7]. In that work Medina summarized the main fuzzy database models in three families (Chapter III): The fuzzy relational model (with a fuzzy degree in each row or tuple), the model based in similarity relations by Buckles and Petry, and the relational models with possibility distributions by Umano, Fukami, Prade, Testemale, Zemankova, Kaendel and other authors. Medina's PhD thesis also embraced the generalizations of fuzzy models. Medina proposed a conceptual framework for fuzzy representation called GEFRED (Generalized Model for Fuzzy Relational Databases) and a language called FSQL (Fuzzy SQL). In the same research group a young mathematician and informatic José Galindo[8] started his PhD research under the supervision of Medina, in order to improve the relational algebra of the GEFRED model, to define a fuzzy relational calculus and to implement other fuzzy comparators. In fact, the possibility and necessity measures, shown by Dubois and Prade, do not only allow the con-

struction of two fuzzy comparators, but 14 of them. The implementation of a new FSQL server running in Oracle and a new GUI interface of the FSQL language was included too.

In these two theses, part of the job was concluded; that is, the physical and logical approaches for development of fuzzy databases. Nevertheless, the conceptual design of fuzzy entities and relations was still missing.

This last step was achieved in 2003 by Angelica Urrutia[9], in her PhD research under the supervision of José Galindo and Mario Piattini. In this work, you find a conceptual fuzzy model, so-called FuzzyEER, and a case tool (FuzzyCASE), to help the database engineers to build the conceptual model for fuzzy databases.

Herein lie the roots of this book, the logical fuzzy models of Medina (1994) and Galindo (1999) on one hand, and, on the other hand, the conceptual fuzzy model of Urrutia (2003).

Personally, I find the name of the book *Fuzzy Databases: Modeling, Design and Implementation* quite right because the work of Galindo, Urrutia, and Piattini is a highly important contribution to understanding the fuzzy database process, not only by professionals of software engineering, but also by computer science students. I hope this book has a real influence in the orientation of the databases courses.

Chapter I, dedicated exclusively to the fuzzy logic, should be appreciated. This chapter could be very useful to new students in this area.

Chapter II brings up to date the classification of fuzzy database models, including some ideas about fuzzy object-oriented database models centered in the relational model, even though these ideas are not used in this book. In spite of this, the contributions of this book will turn out to be very useful for the definition of a complete fuzzy object-oriented database model.

Chapter III is focused on fuzzy database modeling, showing some of the more important approaches by other authors. This chapter is important in order to understanding the importance of the FuzzyEER model defined in Chapter IV, an extension of an EER model to create a model with fuzzy semantics and notations. Although the model has numerous characteristics, the main components of this data modeling tool are: imprecise attributes; fuzzy attributes associated to one or more attributes or with an independent meaning; degrees of fuzzy membership to the model itself, such as fuzzy aggregation, fuzzy entity, weak fuzzy entity, fuzzy relationship; and defined specialization with fuzzy degrees.

Chapter V describes how to represent fuzzy knowledge in relational databases. This methodology is debatable. Nevertheless, as the authors said, it is

complete enough for the immense majority of the applications. On the other hand, the possible lacks in that methodology may be easily solved in each specific application. Chapter VI gives the steps of an algorithm for mapping FuzzyEER models to that methodology. This algorithm relates Chapter IV and V.

Chapter VII describes the more important statements of the FSQL language. This definition improves upon the previous version of this language in many aspects. The educational experience of the authors is noted also in this chapter, which includes a multitude of examples that permits understanding of the utility of each definition.

With all the tools defined in previous chapters, Chapter VIII studies some applications of fuzzy databases. These applications show that fuzzy databases are useful in areas other than management applications (storing and querying information). Of course, FSQL may be used for fuzzy querying, but it can also be used for fuzzy clustering and fuzzy classification, for defining fuzzy dependencies, and for the fuzzy characterization of images in a system of fuzzy image retrieval. The last chapter, the appendices, and references close the book, giving additional information. The open research lines are especially interesting, because they prove that this matter is not closed.

Finally, I borrow the words said by Zadeh in May of 1972, in his Preface of Kaufmann's book, "Professor Kaufmann's treatise is clearly a very important accomplishment. It may a well exert a significant influence on scientific thinking in the years ahead and stimulate much further research on the theory of fuzzy sets and their applications in various field of science and engineering." Well, I think these words match the aim of this book too.

Endnotes

1 See http://www.cs.berkeley.edu/~zadeh

2 This proposal was mentioned by Zadeh, in the he wrote for the preface of the book written by A. Kaufmann in 1977, *Introduction à la théorie des sous-ensembles flous à l'usage des ingénieurs*.

3 The reader can find some ideas related to fuzzy control of engineering systems in the book by Kazuo Tanaka in 1996, *An Introduction to Fuzzy Logic for Practical Applications*.

[4] Dubois, D. (1983). *Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications aux techniques d'aide à la decision.* Doctoral dissertation. Medical and Scientific University, Grenoble, France.

[5] Prade, H. (1982). *Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications au raisonnement naturel.* Doctoral dissertation. Paul Sabatier University, Toulouse, France.

[6] Testemale, C. (1984). *Un système de traitement d'informations incomplètes ou incertaines dans une base de données relationnelles.* Doctoral dissertation. Paul Sabatier University, Toulouse, France.

[7] Medina, J.M. (1994). *Bases de Datos Relacionales Difusas: Modelo teórico y aspecto de su implementación.* Doctoral dissertation. University of Granada, Spain.

[8] Galindo, J. (1999). *Tratamiento de la imprecisión en bases de datos relacionales: Extensión del modelo y adaptación de los SGBD actuales.* Doctoral dissertation. University of Granada, Spain.

[9] Urrutia, A. (2003). *Una definición de un modelo conceptual a una base de datos difusa. Doctoral dissertation.* University of Castilla-La Mancha, Spain.

[10] On February 22, 2005, Dr. Jiménez was awarded with the Trophy Fernand Gallais, who grants the *Ecole Nationale Supérieure des Ingénieurs in Arts Chimiques et Technologiques* of the Polytechnic Institute of Toulouse, France, for his PhD thesis *Gestion des connaissances imparfaites dans les organisations industrielles: Cas d'une industrie manufacturière en Amérique Latine.*

Dr. Leoncio Jiménez[10]

Talca, Chile, June 2005

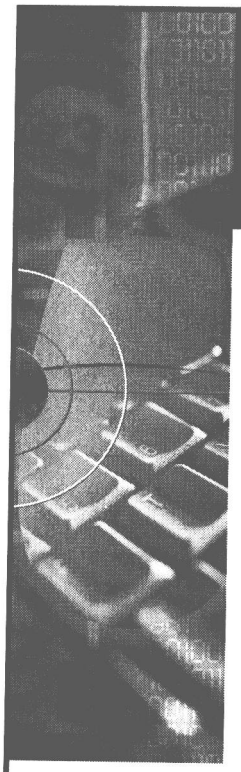
此为试读,需要完整PDF请访问: www.ertongbook.com

Acknowledgments

The authors would like to thank the CyTED project VII-J-RITOS2 (and its main researcher, Professor Nieves R. Brisaboa) and MCYT project TIC2002-00480 for their partial support, which contributed to the final formulation of this work.

In addition, IDBIS Research Group (Intelligent DataBases and Information Systems; see http://idbis.ugr.es or http://frontdb.ugr.es) has been one of our inspirations. The works of these researchers have been and are very important in the field of fuzzy logic and fuzzy databases. Thank you very much.

Of course, many people collaborated directly or indirectly so that this book could finally be published. They know who they are. Thanks to all, and especially to Dra. M. Carmen Aranda for her useful technical revisions in Chapter VIII.



Experience the latest full-text research in the fields of Information Science, Technology & Management

InfoSci-Online

InfoSci-Online is available to libraries to help keep students, faculty and researchers up-to-date with the latest research in the ever-growing field of information science, technology, and management.

The InfoSci-Online collection includes:

- Scholarly and scientific book chapters
- Peer-reviewed journal articles
- Comprehensive teaching cases
- Conference proceeding papers
- All entries have abstracts and citation information
- The full text of every entry is downloadable in .pdf format

Some topics covered:

- Business Management
- Computer Science
- Education Technologies
- Electronic Commerce
- Environmental IS
- Healthcare Information Systems
- Information Systems
- Library Science
- Multimedia Information Systems
- Public Information Systems
- Social Science and Technologies

InfoSci-Online features:

- Easy-to-use
- 6,000+ full-text entries
- Aggregated
- Multi-user access

"...The theoretical bent of many of the titles covered, and the ease of adding chapters to reading lists, makes it particularly good for institutions with strong information science curricula."

— Issues in Science and Technology Librarianship

To receive your free 30-day trial access subscription contact:

Andrew Bundy

Email: abundy@idea-group.com • Phone: 717/533-8845 x29

Web Address: www.infosci-online.com

InfoSci-Online

Full Text • Cutting Edge • Easy Access

A PRODUCT OF IDEA GROUP INC.

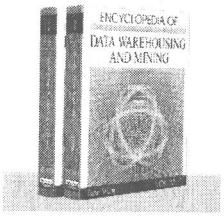
Publishers of Idea Group Publishing, Information Science Publishing, CyberTech Publishing, and IRM Press

infosci-online.com

New Releases from Idea Group Reference

Idea Group REFERENCE

The Premier Reference Source for Information Science and Technology Research



ENCYCLOPEDIA OF
DATA WAREHOUSING AND MINING

Edited by: John Wang,
Montclair State University, USA

Two-Volume Set • April 2005 • 1700 pp
ISBN: 1-59140-557-2; US $495.00 h/c
Pre-Publication Price: US $425.00*
*Pre-pub price is good through one month after the publication date

- Provides a comprehensive, critical and descriptive examination of concepts, issues, trends, and challenges in this rapidly expanding field of data warehousing and mining
- A single source of knowledge and latest discoveries in the field, consisting of more than 350 contributors from 32 countries
- Offers in-depth coverage of evolutions, theories, methodologies, functionalities, and applications of DWM in such interdisciplinary industries as healthcare informatics, artificial intelligence, financial modeling, and applied statistics
- Supplies over 1,300 terms and definitions, and more than 3,200 references

ENCYCLOPEDIA OF
DISTANCE LEARNING

Four-Volume Set • April 2005 • 2500+ pp
ISBN: 1-59140-555-6; US $995.00 h/c
Pre-Pub Price: US $850.00*
*Pre-pub price is good through one month after the publication date

- More than 450 international contributors provide extensive coverage of topics such as workforce training, accessing education, digital divide, and the evolution of distance and online education into a multibillion dollar enterprise
- Offers over 3,000 terms and definitions and more than 6,000 references in the field of distance learning
- Excellent source of comprehensive knowledge and literature on the topic of distance learning programs
- Provides the most comprehensive coverage of the issues, concepts, trends, and technologies of distance learning

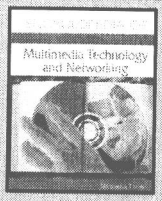
ENCYCLOPEDIA OF
INFORMATION SCIENCE AND TECHNOLOGY
AVAILABLE NOW!

Five-Volume Set • January 2005 • 3807 pp
ISBN: 1-59140-553-X; US $1125.00 h/c

ENCYCLOPEDIA OF
DATABASE TECHNOLOGIES AND APPLICATIONS

April 2005 • 650 pp
ISBN: 1-59140-560-2; US $275.00 h/c
Pre-Publication Price: US $235.00*
*Pre-publication price good through one month after publication date

ENCYCLOPEDIA OF
MULTIMEDIA TECHNOLOGY AND NETWORKING

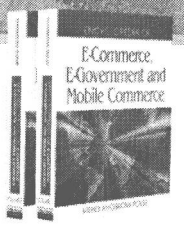


April 2005 • 650 pp
ISBN: 1-59140-561-0; US $275.00 h/c
Pre-Publication Price: US $235.00*
*Pre-pub price is good through one month after publication date

www.idea-group-ref.com

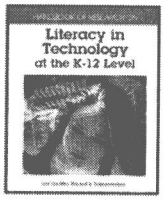
Idea Group Reference is pleased to offer complimentary access to the electronic version for the life of edition when your library purchases a print copy of an encyclopedia

For a complete catalog of our new & upcoming encyclopedias, please contact:
701 E. Chocolate Ave., Suite 200 • Hershey PA 17033, USA • 1-866-342-6657 (toll free) • cust@idea-group.com



ENCYCLOPEDIA OF
E-Commerce, E-Government and Mobile Commerce
Edited by: Mehdi Khosrow-Pour
Two-Volume Set
March 2006; 1200 pp
ISBN: 1-59140-799-0; US $545.00 h/c
Pre-Publication Price: US $460.00*
Online Electronic Access Only**: US $436.00

- Includes quality contributions highlighting current concepts, trends, challenges, applications, and dot.com experiences in the field of e-commerce, e-government, and mobile commerce
- Consists of hundreds of contributions from experts and researchers worldwide, and extensive reference sections which list additional research works



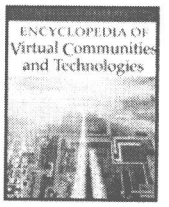
HANDBOOK OF
Research on Literacy in Technology at the K-12 Level
Edited by: Leo Tan Wee Hin & R. Subramaniam
December 2005; 600 pp
ISBN: 1-59140-494-0; US $195.00 h/c
Pre-Publication Price: US $175.00*
Online Electronic Access Only**: US $156.00

- Over 50 international experts have combined their research and practical experience into 35 all-inclusive chapters, redefining the way teaching and learning is dispensed
- This authoritative handbook details the needs of teachers, researchers, and scholars through state-of-the-art perspectives, exposing them to new ideas and interesting developments

Idea Group REFERENCE

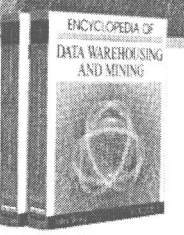
The Premier Reference Source for Information Science and Technology Research

Free online access for the life of the edition with purchase of a print copy.



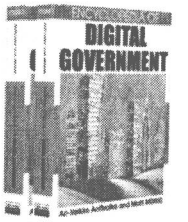
ENCYCLOPEDIA OF
Virtual Communities and Technologies
Edited by: Subhasish Dasgupta
September 2005; 650 pp
ISBN: 1-59140-563-7; US $255.00 h/c
Online Electronic Access Only**: US $204.00

- The hundreds of entries included in this encyclopedia have been authored by 142 leading international researchers, presenting an in-depth analysis of the concepts, technologies, and impacts of virtual environments in our global village
- Offers coverage of upcoming and emerging technologies for virtual communities human-computer interfaces, new networking, mobile computing, web services, and synchronous and asynchronous environments



ENCYCLOPEDIA OF
Data Warehousing and Mining
Edited by: John Wang
Two-Volume Set
July 2005; 1382 pp
ISBN: 1-59140-557-2; US $495.00 h/c
Online Electronic Access Only**: US $396.00

"The Encyclopedia of Data Warehousing and Mining is an accessible yet comprehensive resource for students, researchers and practitioners. The range of topics covered by the encyclopedia is very broad indeed, and it should prove extremely useful as a desktop reference for those active in data mining and warehousing, as well as in related fields."
—Professor Sach Mukherjee, University of Oxford, UK



ENCYCLOPEDIA OF
Digital Government
Edited by: Ari-Veikko Anttiroiko & Matti Malkia
Two-Volume Set
March 2006; 1000+ pp
ISBN: 1-59140-789-3; US $565.00 h/c
Pre-Publication Price: US $480.00*
Online Electronic Access Only**: US $452.00

- Offers the most comprehensive coverage of the issues, concepts, trends, and technologies of digital government
- Containing over 200 detailed articles, this two-volume set provides a broad basis for understanding the issues, theories, and applications faced by public administrations and public organizations as they strive for more effective government through the use of emerging technologies

*Pre-pub price is good through one month after the publication date. **Electronic access is for libraries only and is good for the life of the edition.

701 E. Chocolate Ave., Suite 200 • Hershey PA 17033, USA
1-866-342-6657 (toll free), 717-533-8845 x10, cust@idea-group.com
www.idea-group-ref.com

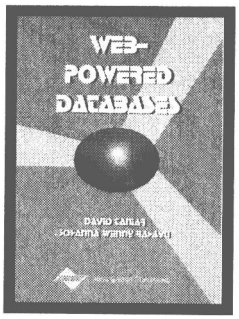
IGP Release!

Web-Powered Databases

David Taniar, PhD
Monash University, Australia

Johanna Wenny Rahayu, PhD
La Trobe University, Australia

Web-Powered Databases provides an excellent snapshot of current research and development activities in the area of Web or Internet databases. Its content supplies potential answers to many questions that have been raised regarding database accesses through the Web. This book also provides a number of case studies of successful Web database applications, including multiple-choice assessment through the Web, an online pay claim, a product catalogue, and content management and dynamic Web pages.



ISBN 1-59140-035-X (h/c)• eISBN 1-59140-092-9• US$89.95 • 340 pages • © 2003

"With the increasing use and development of Internet technology, it makes sense to have a database system implemented on the Web, so that information stored in the database can be more accessible."
–David Taniar, Monash University

It's Easy to Order! Order online at www.idea-group.com or call 717/533-8845 x10!
Mon-Fri 8:30 am-5:00 pm (est) or fax 24 hours a day 717/533-8661

IGP

Idea Group Publishing

Hershey • London • Melbourne • Singapore • Beijing

An excellent addition to your library

Fuzzy Databases: Modeling, Design and Implementation

Table of Contents

Foreword ..... **ix**
Juan Miguel Medina, Spain

Preface ..... **xi**
Leoncio Jiménez, Chile

Chapter I
Introduction to Fuzzy Logic **1**
Fuzzy Sets *2*
Types of Membership Functions *5*
Membership Function Determination *11*
Concepts About Fuzzy Sets *13*
Fuzzy Set Operations *16*
Union and Intersection: t-conorms and t-norms *16*
Complements or Negations *19*
Comparison Operations on Fuzzy Sets *22*
Fuzzy Relations *32*
Operations and Compositions of Fuzzy Relations *33*
Fuzzy Numbers *34*
The Extension Principle *36*
Fuzzy Arithmetic *38*
Possibility Theory *39*
Fuzzy Quantifiers *40*

此为试读，需要完整PDF请访问：www.ertongbook.com